GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 29, 2002, 09:53:06 ; Search time 14.56 Seconds
(without alignments)
15.098 Million cell updates/sec

Title: US-09-734-628-1
Perfect score: 65
Sequence: 1 CDCRGDCFC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 60703

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	65	100.0	9	2	US-08-701-124-3	Sequence 3, Appli
2	65	100.0	9	2	US-08-286-861-16	Sequence 16, Appl
3	65	100.0	9	3	US-09-026-633-1	Sequence 1, Appli
4	65	100.0	9	3	US-09-130-225-3	Sequence 3, Appli
5	65	100.0	9	4	US-09-124-671-33	Sequence 33, Appl
6	65	100.0	9	4	US-09-258-754-211	Sequence 211, App
7	65	100.0	9	4	US-09-139-802-1	Sequence 1, Appli
8	65	100.0	9	4	US-09-042-107-211	Sequence 211, App
9	65	100.0	9	4	US-09-320-424-20	Sequence 20, Appl
10	65	100.0	9	4	US-09-426-680-12	Sequence 12, Appl
11	65	100.0	9	4	US-09-455-061-3	Sequence 3, Appli
12	59	90.8	9	2	US-08-286-861-17	Sequence 17, Appl
13	56	86.2	8	3	US-09-026-633-4	Sequence 4, Appli
14	51	78.5	9	2	US-08-701-124-4	Sequence 4, Appli
15	51	78.5	9	2	US-08-286-861-15	Sequence 15, Appl
16	51	78.5	9	3	US-09-130-225-4	Sequence 4, Appli
17	51	78.5	9	4	US-09-455-061-4	Sequence 4, Appli
18	49	75.4	9	2	US-08-286-861-18	Sequence 18, Appl
19	44	67.7	7	4	US-09-426-680-11	Sequence 11, Appl
20	40	61.5	8	1	US-08-421-702A-22	Sequence 22, Appl
21	40	61.5	8	1	US-08-303-052A-22	Sequence 22, Appl
22	40	61.5	8	1	US-08-421-696A-22	Sequence 22, Appl
23	40	61.5	8	1	US-08-421-697A-22	Sequence 22, Appl
24	40	61.5	8	1	US-08-421-698A-22	Sequence 22, Appl
25	40	61.5	8	2	US-08-421-695A-22	Sequence 22, Appl
26	40	61.5	8	5	PCT-US95-04741-22	Sequence 22, Appl
27	38	58.5	7	2	US-08-286-861-14	Sequence 14, Appl
28	35	53.8	5	1	US-08-212-186A-10	Sequence 10, Appl
29	35	53.8	5	1	US-08-425-238-8	Sequence 8, Appli
30	35	53.8	5	2	US-08-625-695A-10	Sequence 10, Appl
31	35	53.8	5	2	US-08-335-832-42	Sequence 42, Appl
32	35	53.8	5	2	US-08-753-781-35	Sequence 35, Appl
33	35	53.8	5	2	US-08-286-861-37	Sequence 37, Appl
34	35	53.8	5	3	US-09-141-127-15	Sequence 15, Appl
35	35	53.8	5	4	US-08-924-002-10	Sequence 10, Appl
36	35	53.8	6	1	US-08-212-186A-1	Sequence 1, Appli
37	35	53.8	6	1	US-08-212-186A-26	Sequence 26, Appl
38	35	53.8	6	1	US-08-425-238-4	Sequence 4, Appli
39	35	53.8	6	2	US-08-625-695A-1	Sequence 1, Appli
40	35	53.8	6	2	US-08-625-695A-26	Sequence 26, Appl
41	35	53.8	6	2	US-08-286-861-7	Sequence 7, Appli
42	35	53.8	6	4	US-08-924-002-1	Sequence 1, Appli
43	35	53.8	6	4	US-08-924-002-26	Sequence 26, Appl
44	35	53.8	9	1	US-08-421-702A-23	Sequence 23, Appl
45	35	53.8	9	1	US-08-303-052A-23	Sequence 23, Appl

ALIGNMENTS

```
RESULT   1
US-08-701-124-3
; Sequence 3, Application US/08701124
; Patent No. 5846782
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Roelvink, Petrus W.
;    APPLICANT:  Kovesdi, Imre
;    TITLE OF INVENTION:  TARGETING ADENOVIRUS WITH USE OF
;    TITLE OF INVENTION:  CONSTRAINED PEPTIDE MOTIFS
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza - 49th Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/701,124
;      FILING DATE:  21-AUG-1996
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-701-124-3


  Query Match             100.0%;  Score 65;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           |||||||||
Db       1 CDCRGDCFC 9


RESULT   2
US-08-286-861-16
; Sequence 16, Application US/08286861
; Patent No. 5981478
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Ruoslahti, Erkki
;    APPLICANT:  Koivunen, Erkki
;    TITLE OF INVENTION:  No. 5981478el Integrin-Binding Peptides
;    NUMBER OF SEQUENCES:  46
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/286,861
;      FILING DATE:  04-AUG-1994
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/158,001
;      FILING DATE:  24-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-LA 9992
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  circular
US-08-286-861-16


  Query Match           100.0%;  Score 65;  DB 2;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CDCRGDCFC 9
          |||||||||
Db      1 CDCRGDCFC 9


RESULT   3
US-09-026-633-1
; Sequence 1, Application US/09026633
; Patent No. 6025328
; GENERAL INFORMATION:
;  APPLICANT: McMorris, Trevor C.
;  APPLICANT: Kelner, Michael J.
;  TITLE OF INVENTION: Antitumor agents
;  FILE REFERENCE: 103.008US1
;  CURRENT APPLICATION NUMBER: US/09/026,633
;  CURRENT FILING DATE: 1998-02-20
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Amino acid sequence
US-09-026-633-1


  Query Match           100.0%;  Score 65;  DB 3;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CDCRGDCFC 9
          |||||||||
Db      1 CDCRGDCFC 9


RESULT   4
US-09-130-225-3
; Sequence 3, Application US/09130225
; Patent No. 6057155
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Roelvink, Petrus W.
;    APPLICANT:  Kovesdi, Imre
;    TITLE OF INVENTION:  TARGETING ADENOVIRUS WITH USE OF
;    TITLE OF INVENTION:  CONSTRAINED PEPTIDE MOTIFS
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza - 49th Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/130,225
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 8-701124
;      FILING DATE:  21-AUG-1996
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-130-225-3


  Query Match           100.0%;  Score 65;  DB 3;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CDCRGDCFC 9
          |||||||||
Db      1 CDCRGDCFC 9


RESULT   5
US-09-124-671-33
; Sequence 33, Application US/09124671A
; Patent No. 6160088
; GENERAL INFORMATION:
;  APPLICANT: Rothman, James
;  APPLICANT: Mayhew, Mark
;  APPLICANT: Hoe, Mee
;  TITLE OF INVENTION: KDEL RECEPTOR INHIBITORS
;  FILE REFERENCE: 31488
;  CURRENT APPLICATION NUMBER: US/09/124,671A
;  CURRENT FILING DATE: 1998-07-29
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 33
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: alpha-five integrin binding motif
```

US-09-124-671-33

Query Match 100.0%; Score 65; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 CDCRGDCFC 9

RESULT 6
US-09-258-754-211
; Sequence 211, Application US/09258754
; Patent No. 6174687
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; APPLICANT: Rajotte, Daniel
; TITLE OF INVENTION: Methods of Identifying Lung Homing Molecules Using
; TITLE OF INVENTION: Membrane Dipeptidase
; FILE REFERENCE: P-LJ 3443
; CURRENT APPLICATION NUMBER: US/09/258,754
; CURRENT FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 09/042,107
; EARLIER FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 452
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 211
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-258-754-211

Query Match 100.0%; Score 65; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 CDCRGDCFC 9

RESULT 7
US-09-139-802-1
; Sequence 1, Application US/09139802
; Patent No. 6180084
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: NGR Receptor and Methods of Identifying Tumor Homing
; TITLE OF INVENTION: Molecules That Home to Angiogenic Vasculature Using
; TITLE OF INVENTION: Same
; FILE REFERENCE: P-LJ 3203
; CURRENT APPLICATION NUMBER: US/09/139,802
; CURRENT FILING DATE: 1998-08-25
; EARLIER APPLICATION NUMBER: 08/926,914
; EARLIER FILING DATE: 1997-09-10
; EARLIER APPLICATION NUMBER: 08/710,067
; EARLIER FILING DATE: 1996-09-10
; NUMBER OF SEQ ID NOS: 226
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide

US-09-139-802-1

Query Match 100.0%; Score 65; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 CDCRGDCFC 9

RESULT 8
US-09-042-107-211
; Sequence 211, Application US/09042107
; Patent No. 6232287
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Molecules that Home to Various Selected Organs or
; TITLE OF INVENTION: Tissues
; FILE REFERENCE: P-LJ 2892
; CURRENT APPLICATION NUMBER: US/09/042,107
; CURRENT FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 436
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 211
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-042-107-211

Query Match 100.0%; Score 65; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 CDCRGDCFC 9

RESULT 9
US-09-320-424-20
; Sequence 20, Application US/09320424
; Patent No. 6284236
; GENERAL INFORMATION:
; APPLICANT: Wiley, Steven R.
; APPLICANT: Goodwin, Raymond G.
; TITLE OF INVENTION: Cytokine that Induces Apoptosis
; FILE REFERENCE: 2835-E
; CURRENT APPLICATION NUMBER: US/09/320,424
; CURRENT FILING DATE: 1999-05-26
; EARLIER APPLICATION NUMBER: 09/190,046
; EARLIER FILING DATE: 1998-11-10
; EARLIER APPLICATION NUMBER: 09/048,641
; EARLIER FILING DATE: 1998-03-26
; EARLIER APPLICATION NUMBER: 08/670,354
; EARLIER FILING DATE: 1996-06-25
; EARLIER APPLICATION NUMBER: 08/548,368
; EARLIER FILING DATE: 1995-11-01
; EARLIER APPLICATION NUMBER: 08/496,632
; EARLIER FILING DATE: 1995-06-29
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: artificial

```
;   OTHER INFORMATION: peptide
US-09-320-424-20


  Query Match             100.0%;  Score 65;  DB 4;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           |||||||||
Db       1 CDCRGDCFC 9


RESULT  10
US-09-426-680-12
; Sequence 12, Application US/09426680
; Patent No. 6287857
; GENERAL INFORMATION:
;  APPLICANT: Catherine R. O'Riordan
;  APPLICANT: Samuel C. Wadsworth
;  TITLE OF INVENTION: Nucleic Acid Delivery Vehicles
;  FILE REFERENCE: GA0103USB2
;  CURRENT APPLICATION NUMBER: US/09/426,680
;  CURRENT FILING DATE: 1999-10-25
;  EARLIER APPLICATION NUMBER: PCT/US99/02680
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 12
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: human
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (0)...(0)
US-09-426-680-12


  Query Match             100.0%;  Score 65;  DB 4;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           |||||||||
Db       1 CDCRGDCFC 9


RESULT  11
US-09-455-061-3
; Sequence 3, Application US/09455061
; Patent No. 6329190
; GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Roelvink, Petrus W.
;    APPLICANT:  Kovesdi, Imre
;    TITLE OF INVENTION:  TARGETING ADENOVIRUS WITH USE OF
;    TITLE OF INVENTION:  CONSTRAINED PEPTIDE MOTIFS
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza - 49th Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/455,061
;      FILING DATE:  06-DEC-1999
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 9-130225
;      FILING DATE:  06-AUG-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 8-701124
;      FILING DATE:  21-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hefner, M. Daniel
;      REGISTRATION NUMBER:  41,826
;      REFERENCE/DOCKET NUMBER:  203128
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-455-061-3


  Query Match             100.0%;  Score 65;  DB 4;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           |||||||||
Db       1 CDCRGDCFC 9


RESULT  12
US-08-286-861-17
; Sequence 17, Application US/08286861
; Patent No. 5981478
;  GENERAL INFORMATION:
;    APPLICANT:  Ruoslahti, Erkki
;    APPLICANT:  Koivunen, Erkki
;    TITLE OF INVENTION:  No. 5981478el Integrin-Binding Peptides
;    NUMBER OF SEQUENCES:  46
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/286,861
;      FILING DATE:  04-AUG-1994
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/158,001
;      FILING DATE:  24-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-LA 9992
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  circular
US-08-286-861-17


  Query Match             90.8%;  Score 59;  DB 2;  Length 9;
  Best Local Similarity   88.9%;  Pred. No. 1.7e+05;
```

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 CDCRGDCFC 9
          ||||||| |
Db      1 CDCRGDCLC 9
```

```
RESULT 13
US-09-026-633-4
; Sequence 4, Application US/09026633
; Patent No. 6025328
; GENERAL INFORMATION:
;  APPLICANT: McMorris, Trevor C.
;  APPLICANT: Kelner, Michael J.
;  TITLE OF INVENTION: Antitumor agents
;  FILE REFERENCE: 103.008US1
;  CURRENT APPLICATION NUMBER: US/09/026,633
;  CURRENT FILING DATE: 1998-02-20
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;  LENGTH: 8
;  TYPE: PRT
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Amino acid sequence
US-09-026-633-4
```

Query Match 86.2%; Score 56; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      2 DCRGDCFC 9
          ||||||||
Db      1 DCRGDCFC 8
```

```
RESULT 14
US-08-701-124-4
; Sequence 4, Application US/08701124
; Patent No. 5846782
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Roelvink, Petrus W.
;    APPLICANT:  Kovesdi, Imre
;    TITLE OF INVENTION:  TARGETING ADENOVIRUS WITH USE OF
;    TITLE OF INVENTION:  CONSTRAINED PEPTIDE MOTIFS
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza - 49th Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/701,124
;      FILING DATE:  21-AUG-1996
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-701-124-4
```

Query Match 78.5%; Score 51; DB 2; Length 9;
Best Local Similarity 77.8%; Pred. No. 1.7e+05;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 CDCRGDCFC 9
          | ||||| |
Db      1 CXCRGDCXC 9
```

```
RESULT 15
US-08-286-861-15
; Sequence 15, Application US/08286861
; Patent No. 5981478
;  GENERAL INFORMATION:
;    APPLICANT:  Ruoslahti, Erkki
;    APPLICANT:  Koivunen, Erkki
;    TITLE OF INVENTION:  No. 5981478el Integrin-Binding Peptides
;    NUMBER OF SEQUENCES:  46
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/286,861
;      FILING DATE:  04-AUG-1994
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/158,001
;      FILING DATE:  24-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-LA 9992
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  circular
US-08-286-861-15
```

Query Match 78.5%; Score 51; DB 2; Length 9;
Best Local Similarity 77.8%; Pred. No. 1.7e+05;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 CDCRGDCFC 9
          | ||||| |
Db      1 CXCRGDCXC 9
```

Search completed: May 29, 2002, 09:56:53
Job time: 227 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 29, 2002, 09:42:45 ; Search time 20.02 Seconds
(without alignments)
43.197 Million cell updates/sec

Title: US-09-734-628-1
Perfect score: 65
Sequence: 1 CDCRGDCFC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being prin and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	44	67.7	577	2	B37057	integrin beta-6 ch
2	44	67.7	788	2	A37057	integrin beta-6 ch
3	43	66.2	1076	2	T26044	hypothetical prote
4	43	66.2	1627	2	S65464	pregnancy-associat
5	43	66.2	4753	1	A47437	LDL-receptor-relat
6	42.5	65.4	48	2	S29216	neurotoxin Tx2 - s
7	42.5	65.4	49	2	S29215	neurotoxin Tx2 - s
8	42.5	65.4	53	2	S29214	neurotoxin Tx2 - s
9	41	63.1	69	2	A55011	metallothionein-li
10	41	63.1	458	2	A84306	hypothetical prote
11	41	63.1	736	2	T06757	hypothetical prote
12	41	63.1	3672	2	T23433	hypothetical prote
13	41	63.1	3704	2	T37316	probable laminin a
14	40	61.5	195	1	TVVPA	small T antigen -
15	40	61.5	195	2	S22562	small T antigen -
16	40	61.5	313	2	S44208	extracellular matr
17	40	61.5	421	1	TVVPMP	middle T antigen -
18	40	61.5	421	2	S22561	middle T antigen -
19	40	61.5	440	1	TVVPM	middle T antigen -
20	39.5	60.8	246	2	A24609	acidic epididymal
21	39	60.0	30	2	JX0057	trypsin inhibitor
22	39	60.0	32	2	A05076	metallothionein -
23	39	60.0	50	2	T38209	probable metalloth
24	39	60.0	60	2	S30567	metallothionein -
25	39	60.0	60	2	JC2420	metallothionein -
26	39	60.0	61	1	SMBO2	metallothionein II
27	39	60.0	61	2	S00808	metallothionein Ia
28	39	60.0	61	2	S00810	metallothionein Ic
29	39	60.0	61	2	S00809	metallothionein Ib
30	39	60.0	61	2	I46602	metallothionein -
31	39	60.0	68	2	S44392	metallothionein 3
32	39	60.0	656	2	JC2005	integrin beta-5 ch
33	39	60.0	799	2	A38308	integrin beta-5 ch
34	39	60.0	850	2	S56015	gastric mucin MUC5
35	39	60.0	1182	2	I48378	hairless protein -
36	39	60.0	1291	2	T21694	hypothetical prote
37	39	60.0	1321	2	JE0352	mucin MUC5B, trach
38	39	60.0	1373	2	JE0095	gastric mucin MUC5
39	39	60.0	1513	2	A54895	mucin 2, intestina
40	39	60.0	3020	2	A43932	mucin 2 precursor,
41	38.5	59.2	788	2	I51530	integrin beta-3 su
42	38	58.5	40	1	SMFF	metallothionein Mt
43	38	58.5	40	2	B61194	metallothionein Mt
44	38	58.5	60	2	S31723	metallothionein -
45	38	58.5	60	2	B27490	metallothionein B

ALIGNMENTS

RESULT 1
B37057
integrin beta-6 chain - guinea pig (fragment)
C;Species: Cavia porcellus (guinea pig)
5-Feb-1991 #sequence_revision 13-Sep-1991 #text_change 20-Aug-1999
on: B37057
d, D.; Rozzo, C.; Starr, L.; Quaranta, V.; Erle, D.J.; Pytela, R.
Chem. 265, 11502-11507, 1990
Complete amino acid sequence of a novel integrin beta subunit (beta6) identi
e number: A37057; MUID:90307659
n: B37057
preliminary
type: mRNA
1-577 <SHE>
erences: GB:M35197; GB:J05522; NID:g191277; PIDN:AAA37043.1; PID:g553845
authors translated the codon AAA for residue 88 as Asn, AAC for residue 9
for residue 355 as Met, GAG for residue 363 as Thr, ACC for residue 364 a

ly: integrin beta chain; laminin-type EGF-like homology
C;Keywords: cell adhesion; cytoskeleton; transmembrane protein

Query Match 67.7%; Score 44; DB 2; Length 577;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
| |||:|
Db 370 CSGRGDCYC 378

RESULT 2
A37057
integrin beta-6 chain - human
C;Species: Homo sapiens (man)
C;Date: 15-Feb-1991 #sequence_revision 13-Sep-1991 #text_change 19-Jan-2001
C;Accession: A37057; I69201
R;Sheppard, D.; Rozzo, C.; Starr, L.; Quaranta, V.; Erle, D.J.; Pytela, R.
J. Biol. Chem. 265, 11502-11507, 1990
A;Title: Complete amino acid sequence of a novel integrin beta subunit (beta6) identi
A;Reference number: A37057; MUID:90307659
A;Accession: A37057
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-788 <SHE>
A;Cross-references: GB:M35198; GB:J05522; NID:g186506; PIDN:AAA36122.1; PID:g186507
R;Jiang, W.M.; Jenkins, D.; Yuan, Q.; Leung, E.; Choo, K.H.; Watson, J.D.; Krissansen
Int. Immunol. 4, 1031-1040, 1992
A;Title: The gene organization of the human beta 7 subunit, the common beta subunit o
A;Reference number: I54749; MUID:93002753
A;Accession: I69201
A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA
A;Residues: 116-157,'R',159-197 <JIA>
A;Cross-references: GB:S49380; NID:g257588; PIDN:AAB23690.1; PID:g257589
C;Genetics:
A;Gene: GDB:ITGB6
A;Cross-references: GDB:131392; OMIM:147558
A;Map position: 2pter-2qter
C;Superfamily: integrin beta chain; laminin-type EGF-like homology
C;Keywords: blocked amino end; cell adhesion; cytoskeleton; glycoprotein; lipoprotein; m
F;708-730/Domain: transmembrane #status predicted <TRM>
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F;7/Binding site: palmitate (Cys) (covalent) #status predicted
F;16,48,97,260,387,396,463,471/Binding site: carbohydrate (Asn) (covalent) #status predi

Query Match 67.7%; Score 44; DB 2; Length 788;
Best Local Similarity 66.7%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
| |||:|
511 CSGRGDCYC 519

RESULT 3
T26044
hypothetical protein W01C8.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T26044
R;Nhan, M.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid W01C8.
A;Reference number: Z20142
A;Accession: T26044
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1076 <NHA>
A;Cross-references: EMBL:U41508; PIDN:AAA82623.1; CESP:W01C8.3
C;Genetics:
A;Gene: CESP:W01C8.3
A;Introns: 59/3; 92/2; 157/3; 189/3; 220/2; 251/3; 275/2; 319/1; 374/3; 407/2

Query Match 66.2%; Score 43; DB 2; Length 1076;
Best Local Similarity 85.7%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CDCRGDC 7
| |||||
Db 276 CGCRGDC 282

RESULT 4
S65464
pregnancy-associated plasma protein A precursor - human
N;Alternate names: PAPP-A
C;Species: Homo sapiens (man)
C;Date: 22-Nov-1996 #sequence_revision 22-Nov-1996 #text_change 05-Nov-1999
C;Accession: S65464; S65463; A54220; I38097
R;Haaning, J.; Oxvig, C.; Overgaard, M.T.; Ebbesen, P.; Kristensen, T.; Sottrup-Jensen,
submitted to the EMBL Data Library, June 1995
A;Description: Complete cDNA sequence of the preproform of human pregnancy-associated pl
A;Reference number: S65464
A;Accession: S65464
A;Molecule type: mRNA
A;Residues: 1-1627 <HAA>
A;Cross-references: EMBL:U28727; NID:g1142969; PIDN:AAC50543.1; PID:g1142970
R;Haaning, J.; Oxvig, C.; Overgaard, M.T.; Ebbesen, P.; Kristensen, T.; Sottrup-Jensen,
Eur. J. Biochem. 237, 159-163, 1996
A;Title: Complete cDNA sequence of the preproform of human pregnancy-associated plasma p
A;Reference number: S65463; MUID:96203921
A;Accession: S65463
A;Molecule type: mRNA
A;Residues: 1-102 <HAW>
A;Cross-references: EMBL:U28727
A;Note: the authors translated the codon CGA for residue 101 as Thr
R;Kristensen, T.; Oxvig, C.; Sand, O.; Moller, N.P.H.; Sottrup-Jensen, L.
Biochemistry 33, 1592-1598, 1994
A;Title: Amino acid sequence of human pregnancy-associated plasma protein-A derived f
A;Reference number: A54220; MUID:94146014
A;Accession: A54220
A;Molecule type: mRNA
A;Residues: 77-1627 <KRI>
A;Cross-references: GB:X68280; NID:g394649; PIDN:CAA48341.1; PID:g394650
R;Oxvig, C.; Sand, O.; Kristensen, T.; Gleich, G.J.; Sottrup-Jensen, L.
J. Biol. Chem. 268, 12243-12246, 1993
A;Title: Circulating human pregnancy-associated plasma protein-A is disulfide-bridged
A;Reference number: I38097; MUID:93286045
A;Accession: I38097
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 77-1627 <RES>
A;Cross-references: EMBL:X68280; NID:g394649; PIDN:CAA48341.1; PID:g394650
C;Genetics:
A;Gene: GDB:PAPPA
A;Cross-references: GDB:134729; OMIM:176385
A;Map position: 9q33.1-9q33.1
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-80/Domain: propeptide #status predicted <PRO>
F;81-1627/Product: pregnancy-associated plasma protein A #status predicted <MAT>

Query Match 66.2%; Score 43; DB 2; Length 1627;
Best Local Similarity 66.7%; Pred. No. 84;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|| :||| |
Db 1600 CDLQGDCAC 1608

RESULT 5
A47437
LDL-receptor-related protein - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 18-Aug-2000
C;Accession: A47437; S27801; T21547
R;Yochem, J.; Greenwald, I.
Proc. Natl. Acad. Sci. U.S.A. 90, 4572-4576, 1993
A;Title: A gene for a low density lipoprotein receptor-related protein in the nematod
A;Reference number: A47437; MUID:93281621
A;Accession: A47437
A;Molecule type: DNA
A;Residues: 1-4753 <YOC>
A;Cross-references: GB:M96150; NID:g156359; PIDN:AAA28105.1; PID:g156360
A;Note: nucleotide sequence not given; translation not complete in this paper
R;Yochem, J.; Greenwald, I.
submitted to the EMBL Data Library, July 1992
A;Description: A gene for an LDL receptor-related protein (LPR) in the nematode C.ele
A;Reference number: S27801
A;Accession: S27801
A;Molecule type: DNA
A;Residues: 1-4753 <YO2>
A;Cross-references: EMBL:M96150; NID:g156359; PIDN:AAA28105.1; PID:g156360
R;Wilkinson, J.
submitted to the EMBL Data Library, June 1996
A;Reference number: Z19439
A;Accession: T21547
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-4753 <WIL>
A;Cross-references: EMBL:Z73907; PIDN:CAA98124.1; GSPDB:GN00019; CESP:F29D11.1
A;Experimental source: clone F29D11
C;Genetics:
A;Gene: LPR

```
A;Map position: 1
A;Introns: 31/1; 88/1; 132/1; 172/3; 219/1; 298/1; 463/2; 526/2; 585/3; 780/2; 874/2; 97
15/1
C;Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding
C;Keywords: tandem repeat; transmembrane protein
F;53-87/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F;92-131/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F;138-175/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F;182-218/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F;223-257/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F;262-297/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F;302-336/Domain: EGF homology <EGF1>
F;1054-1095/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F;1101-1138/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F;1146-1182/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F;1187-1223/Domain: LDL receptor ligand-binding repeat homology <LD10>
F;1228-1263/Domain: LDL receptor ligand-binding repeat homology <LD11>
F;1270-1307/Domain: LDL receptor ligand-binding repeat homology <LD12>
F;1313-1350/Domain: LDL receptor ligand-binding repeat homology <LD13>
F;1359-1396/Domain: LDL receptor ligand-binding repeat homology <LD14>
F;1441-1475/Domain: EGF homology <EGF>
F;1611-1654/Domain: LDL receptor YWTD-containing repeat homology <YW33>
F;2792-2829/Domain: LDL receptor ligand-binding repeat homology <LD15>
F;2834-2868/Domain: LDL receptor ligand-binding repeat homology <LD16>
F;2874-2912/Domain: LDL receptor ligand-binding repeat homology <LD17>
F;2919-2956/Domain: LDL receptor ligand-binding repeat homology <LD18>
F;2961-2997/Domain: LDL receptor ligand-binding repeat homology <LD19>
F;3006-3044/Domain: LDL receptor ligand-binding repeat homology <LD20>
F;3049-3093/Domain: LDL receptor ligand-binding repeat homology <LD21>
F;3100-3135/Domain: LDL receptor ligand-binding repeat homology <LD22>
F;3140-3174/Domain: LDL receptor ligand-binding repeat homology <LD23>
F;3187-3222/Domain: LDL receptor ligand-binding repeat homology <LD24>
F;3586-3623/Domain: EGF homology <EGX1>
F;3627-3666/Domain: LDL receptor ligand-binding repeat homology <LD25>
F;3671-3705/Domain: LDL receptor ligand-binding repeat homology <LD26>
F;3709-3746/Domain: LDL receptor ligand-binding repeat homology <LD27>
F;3753-3788/Domain: LDL receptor ligand-binding repeat homology <LD28>
F;3793-3830/Domain: LDL receptor ligand-binding repeat homology <LD29>
F;3833-3871/Domain: LDL receptor ligand-binding repeat homology <LD30>
F;3878-3912/Domain: LDL receptor ligand-binding repeat homology <LD31>
F;3917-3951/Domain: LDL receptor ligand-binding repeat homology <LD32>
F;3959-3995/Domain: LDL receptor ligand-binding repeat homology <LD33>
F;4000-4040/Domain: LDL receptor ligand-binding repeat homology <LD34>
F;4049-4083/Domain: LDL receptor ligand-binding repeat homology <LD35>
F;4092-4130/Domain: EGF homology <EGF2>
F;4343-4386/Domain: LDL receptor YWTD-containing repeat homology <YW38>

  Query Match             66.2%;  Score 43;  DB 1;  Length 4753;
  Best Local Similarity   75.0%;  Pred. No. 1.9e+02;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCF 8
           | | ||||
Db     361 CSCLGDCF 368


RESULT   6
S29216
neurotoxin Tx2 - spider (Phoneutria nigriventer)
C;Species: Phoneutria nigriventer
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999
C;Accession: S29216
R;do Nascimento Cordeiro, M.; Ribeiro Diniz, C.; do Carmo Valentim, A.; von Eickstedt, V
FEBS Lett. 310, 153-156, 1992
A;Title: The purification and amino acid sequences of four Tx2 neurotoxins from the venc
A;Reference number: S29214; MUID:93011905
A;Accession: S29216
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-48 <COR>
C;Superfamily: curtatoxin

  Query Match             65.4%;  Score 42.5;  DB 2;  Length 48;
  Best Local Similarity   58.3%;  Pred. No. 7.1;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    3;  Gaps    1;

Qy       1 CDC---RGDCFC 9
           |||   ||:| |
Db      14 CDCCGERGECVC 25


RESULT   7
S29215
neurotoxin Tx2 - spider (Phoneutria nigriventer)
C;Species: Phoneutria nigriventer
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 15-Oct-1999
C;Accession: S29215; B39305
R;do Nascimento Cordeiro, M.; Ribeiro Diniz, C.; do Carmo Valentim, A.; von Eickstedt
FEBS Lett. 310, 153-156, 1992
A;Title: The purification and amino acid sequences of four Tx2 neurotoxins from the v
A;Reference number: S29214; MUID:93011905
A;Accession: S29215
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-49 <COR>
R;Rezende Jr., L.; Cordeiro, M.N.; Oliveira, E.B.; Diniz, C.R.
Toxicon 29, 1225-1233, 1991
A;Title: Isolation of neurotoxic peptides from the venom of the 'armed' spider Phoneu
A;Reference number: A39305; MUID:92196803
A;Accession: B39305
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-11 <REZ>
C;Superfamily: curtatoxin
C;Keywords: neurotoxin; venom

  Query Match             65.4%;  Score 42.5;  DB 2;  Length 49;
  Best Local Similarity   58.3%;  Pred. No. 7.3;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    3;  Gaps    1;

Qy       1 CDC---RGDCFC 9
           |||   ||:| |
Db      14 CDCCGERGECVC 25


RESULT   8
S29214
neurotoxin Tx2 - spider (Phoneutria nigriventer)
C;Species: Phoneutria nigriventer
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999
C;Accession: S29214
R;do Nascimento Cordeiro, M.; Ribeiro Diniz, C.; do Carmo Valentim, A.; von Eickstedt
FEBS Lett. 310, 153-156, 1992
A;Title: The purification and amino acid sequences of four Tx2 neurotoxins from the v
A;Reference number: S29214; MUID:93011905
A;Accession: S29214
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-53 <COR>
C;Superfamily: curtatoxin

  Query Match             65.4%;  Score 42.5;  DB 2;  Length 53;
  Best Local Similarity   58.3%;  Pred. No. 7.7;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    3;  Gaps    1;

Qy       1 CDC---RGDCFC 9
           |||   ||:| |
Db      14 CDCCGERGECVC 25


RESULT   9
A55011
```

metallothionein-like protein YOR031w - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein O2675
C;Species: Saccharomyces cerevisiae
C;Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 21-Jul-2000
C;Accession: A55011; S66897
R;Culotta, V.C.; Howard, W.R.; Liu, X.F.
J. Biol. Chem. 269, 25295-25302, 1994
A;Title: CRS5 encodes a metallothionein-like protein in Saccharomyces cerevisiae.
A;Reference number: A55011; MUID:95014318
A;Accession: A55011
A;Molecule type: DNA
A;Residues: 1-69 <CUL>
A;Cross-references: GB:L29056; NID:g499891; PIDN:AAA66061.1; PID:g499892
R;de Haan, M.; Grivell, L.A.; Maarse, A.C.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S66877
A;Accession: S66897
A;Molecule type: DNA
A;Residues: 1-8 <DEH>
A;Cross-references: EMBL:Z74939; MIPS:YOR031w
A;Experimental source: strain S288C
A;Note: in strain S288C YOR031w is a pseudogene with an inframe stopcodon
C;Genetics:
A;Gene: CRS5
A;Map position: 15R
A;Note: YOR031w
C;Function:
A;Description: involved in copper homeostasis and detoxification

Query Match 63.1%; Score 41; DB 2; Length 69;
Best Local Similarity 71.4%; Pred. No. 15;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CDCRGDC 7
||| |:|
Db 6 CDCEGEC 12

RESULT 10
A84306
hypothetical protein Vng1524c [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: A84306
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483
A;Accession: A84306
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-458 <STO>
A;Cross-references: GB:AE004437; NID:g10581011; PIDN:AAG19813.1; GSPDB:GN00138
C;Genetics:
A;Gene: VNG1524C
C;Superfamily: ornithine--oxo-acid aminotransferase

Query Match 63.1%; Score 41; DB 2; Length 458;
Best Local Similarity 55.6%; Pred. No. 63;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
| | |:| |
Db 198 CTCEGECSC 206

RESULT 11
T06757
hypothetical protein F15B8.180 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 22-Oct-1999
C;Accession: T06757
R;Quetier, F.; Benes, V.; Rechmann, S.; Borkova, D.; Ansorge, W.; Salanoubat, M.; Mew
submitted to the Protein Sequence Database, April 1999
A;Reference number: Z15794
A;Accession: T06757
A;Molecule type: DNA
A;Residues: 1-736 <QUE>
A;Cross-references: EMBL:AL049660; GSPDB:GN00061; ATSP:F15B8.180
A;Experimental source: cultivar Columbia; BAC clone F15B8
C;Genetics:
A;Gene: ATSP:F15B8.180
A;Map position: 3
A;Introns: 114/3; 146/1; 208/2; 293/3; 365/3; 384/3; 429/3; 467/3; 536/2; 563/2; 640/

Query Match 63.1%; Score 41; DB 2; Length 736;
Best Local Similarity 53.8%; Pred. No. 90;
Matches 7; Conservative 1; Mismatches 1; Indels 4; Gaps 1;

Qy 1 CDCRGDC----FC 9
|||: || ||
Db 257 CDCKYDCLWGRFC 269

RESULT 12
T23433
hypothetical protein K08C7.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C;Accession: T23433
R;Berks, M.
submitted to the EMBL Data Library, March 1996
A;Reference number: Z19740
A;Accession: T23433
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-3672 <WIL>
A;Cross-references: EMBL:Z70286; PIDN:CAA94293.1; GSPDB:GN00022; CESP:K08C7.3
A;Experimental source: clone K08C7
C;Genetics:
A;Gene: CESP:K08C7.3
A;Map position: 4
A;Introns: 66/1; 284/3; 563/1; 1187/3; 1248/3; 1300/1; 1460/1; 1623/3; 2361/3; 2988/3
C;Superfamily: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGF-lik

Query Match 63.1%; Score 41; DB 2; Length 3672;
Best Local Similarity 55.6%; Pred. No. 3e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|| | |:|
Db 668 CDSNGQCYC 676

RESULT 13
T37316
probable laminin alpha chain - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000
C;Accession: T37316
R;Joh, K.; Zhu, K.; Hedgecock, E.M.; Inoue, T.; Hori, K.
submitted to the EMBL Data Library, August 1998
A;Description: Laminin alpha chain gene in the nematode C. elegans.
A;Reference number: Z21681
A;Accession: T37316
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-3704 <JOH>
A;Cross-references: EMBL:AB016806; PIDN:BAA32347.1

A;Experimental source: strain N2
C;Genetics:
A;Gene: epi-1
A;Map position: IV
A;Introns: 66/1; 284/3; 563/1; 1187/3; 1248/3; 1300/1; 1460/1; 1623/3; 2361/3; 2988/3; 3
C;Superfamily: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGF-like h

Query Match 63.1%; Score 41; DB 2; Length 3704;
Best Local Similarity 55.6%; Pred. No. 3e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|| | |:|
Db 668 CDSNGQCYC 676

RESULT 14
TVVPA
small T antigen - mouse polyomavirus
C;Species: Polyomavirus muris (mouse polyomavirus)
C;Date: 31-Jul-1980 #sequence_revision 31-Jul-1980 #text_change 24-Sep-1999
C;Accession: C03635; B36761; C28838; A03614
R;Soeda, E.; Arrand, J.R.; Smolar, N.; Walsh, J.E.; Griffin, B.E.
Nature 283, 445-453, 1980
A;Title: Coding potential and regulatory signals of the polyoma virus genome.
A;Reference number: A03635; MUID:80099647
A;Accession: C03635
A;Molecule type: DNA
A;Residues: 1-195 <SOE>
A;Cross-references: GB:J02288; GB:J02290; GB:J02291; GB:J02292; GB:K00932; GB:K00997; GB
A;Experimental source: strain A2
R;Friedmann, T.; Esty, A.; LaPorte, P.; Deininger, P.
Cell 17, 715-724, 1979
A;Title: The nucleotide sequence and genome organization of the polyoma early region: ex
A;Reference number: A36761; MUID:80001963
A;Accession: B36761
A;Molecule type: DNA
A;Residues: 1-195 <FRI>
A;Cross-references: GB:J02288; GB:J02290; GB:J02291; GB:J02292; GB:K00932; GB:K00997; GB
A;Experimental source: strain 3
R;Rothwell, V.M.; Folk, W.R.
J. Virol. 48, 472-480, 1983
A;Title: Comparison of the DNA sequence of the Crawford small-plaque variant of polyomav
A;Reference number: A28838; MUID:84011043
A;Accession: C28838
A;Molecule type: DNA
A;Residues: 1-195 <ROT>
A;Cross-references: GB:K02737; NID:g332788
A;Experimental source: strain Crawford small-plaque
A;Note: this ORF is not annotated in GenBank entry PLYCSP
C;Genetics:
A;Introns: 192/1
C;Superfamily: small T antigen; dnaJ amino-terminal homology
C;Keywords: early protein
F;12-62/Domain: dnaJ amino-terminal homology #status atypical <DNJ>

Query Match 61.5%; Score 40; DB 1; Length 195;
Best Local Similarity 53.8%; Pred. No. 47;
Matches 7; Conservative 1; Mismatches 1; Indels 4; Gaps 1;

Qy 1 CDCR----GDCFC 9
|| | |:|||
Db 138 CDARCLVLGECFC 150

RESULT 15
S22562
small T antigen - mouse plasmid L factor
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 26-Aug-1999
C;Accession: S22562
R;Yoshimura, H.; Ikeda, Y.; Yoshimoto, M.; Tamaki, S.; Hanada, K.; Kusano, T.; Kohda,
Nucleic Acids Res. 19, 3633-3639, 1991
A;Title: Structural and functional analysis of a polyoma-related mammalian plasmid (L
A;Reference number: S22560; MUID:91305109
A;Accession: S22562
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-195 <YOS>
A;Cross-references: EMBL:X59849; NID:g52899; PIDN:CAA42512.1; PID:g52902
C;Genetics:
A;Genome: plasmid
A;Introns: 192/1
C;Superfamily: small T antigen; dnaJ amino-terminal homology
F;12-62/Domain: dnaJ amino-terminal homology #status atypical <DNJ>

Query Match 61.5%; Score 40; DB 2; Length 195;
Best Local Similarity 53.8%; Pred. No. 47;
Matches 7; Conservative 1; Mismatches 1; Indels 4; Gaps 1;

Qy 1 CDCR----GDCFC 9
|| | |:|||
Db 138 CDARCLVLGECFC 150

Search completed: May 29, 2002, 09:52:37
Job time: 592 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 29, 2002, 09:52:41 ; Search time 9.93 Seconds
(without alignments)
35.093 Million cell updates/sec

Title: US-09-734-628-1
Perfect score: 65
Sequence: 1 CDCRGDCFC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	44	67.7	577	1	ITB6_CAVPO	P18563 cavia porce
2	44	67.7	787	1	ITB6_MOUSE	Q9z0t9 mus musculu
3	44	67.7	788	1	ITB6_HUMAN	P18564 homo sapien
4	43	66.2	1627	1	PAPA_HUMAN	Q13219 homo sapien
5	43	66.2	4753	1	LRP_CAEEL	Q04833 caenorhabdi
6	42.5	65.4	49	1	TX25_PHONI	P29424 phoneutria
7	42.5	65.4	82	1	TX26_PHONI	P29425 phoneutria
8	42.5	65.4	82	1	TX5A_PHONI	O76199 phoneutria
9	42.5	65.4	88	1	TX21_PHONI	P29423 phoneutria
10	42.5	65.4	115	1	TX1A_PHONI	O76198 phoneutria
11	41	63.1	69	1	CRS5_YEAST	P41902 saccharomyc
12	41	63.1	3672	1	LML2_CAEEL	Q21313 caenorhabdi
13	40.5	62.3	799	1	ITBN_DROME	Q27591 drosophila
14	40	61.5	195	1	TASM_POVMA	P03078 mouse polyo
15	40	61.5	421	1	TAMI_POVMA	P03077 mouse polyo
16	40	61.5	421	1	TAMI_POVMC	P12906 mouse polyo
17	40	61.5	440	1	TAMI_POVM3	P03076 mouse polyo
18	39.5	60.8	246	1	AEG_RAT	P12020 rattus norv
19	39	60.0	30	1	ITR1_MOMCH	P10294 momordica c
20	39	60.0	30	1	ITR3_MOMCO	P82410 momordica c
21	39	60.0	60	1	MTA_CHAAC	O93593 chaenocepha
22	39	60.0	60	1	MTA_CHIHA	O13258 chionodraco
23	39	60.0	60	1	MTA_NOTCO	O73914 notothenia
24	39	60.0	60	1	MTA_PAGBE	O93609 pagothenia
25	39	60.0	60	1	MTA_SPAAU	P52727 sparus aura
26	39	60.0	60	1	MTB_CHAAC	P52724 chaenocepha
27	39	60.0	60	1	MTB_CHIHA	O13259 chionodraco
28	39	60.0	60	1	MTB_DICLA	Q9ptg9 dicentrarch
29	39	60.0	60	1	MTB_PAGBE	Q92145 pagothenia
30	39	60.0	60	1	MT_OREMO	P52726 oreochromis
31	39	60.0	60	1	MT_PAGMA	Q91b50 pagrus majo
32	39	60.0	60	1	MT_PARCR	O93450 parachaenic
33	39	60.0	60	1	MT_PERFL	P52725 perca fluvi
34	39	60.0	60	1	MT_PLEPL	P07216 pleuronecte
35	39	60.0	60	1	MT_PSEAM	P55945 pseudopleur
36	39	60.0	61	1	MT1A_BOVIN	P04356 bos taurus
37	39	60.0	61	1	MT1A_PIG	P49068 sus scrofa
38	39	60.0	61	1	MT1B_SHEEP	P09577 ovis aries
39	39	60.0	61	1	MT1C_PIG	P79376 sus scrofa
40	39	60.0	61	1	MT1C_SHEEP	P09578 ovis aries
41	39	60.0	61	1	MT1E_PIG	P79431 sus scrofa
42	39	60.0	65	1	MT_PARLI	P80367 paracentrot
43	39	60.0	68	1	MT3_HORSE	P37360 equus cabal
44	39	60.0	655	1	ITB5_PAPCY	Q07441 papio cynoc
45	39	60.0	799	1	ITB5_HUMAN	P18084 homo sapien

ALIGNMENTS

RESULT 1
ITB6_CAVPO
ID ITB6_CAVPO STANDARD; PRT; 577 AA.
AC P18563;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Integrin beta-6 (Fragment).
GN ITGB6.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HARTLEY;
RX MEDLINE=90307659; PubMed=2365683;
RA Sheppard D., Rozzo C., Starr L., Quaranta V., Erle D.J., Pytela R.;
RT "Complete amino acid sequence of a novel integrin beta subunit (beta
RT 6) identified in epithelial cells using the polymerase chain
RT reaction.";
RL J. Biol. Chem. 265:11502-11507(1990).
CC -!- FUNCTION: INTEGRIN ALPHA-V/BETA-6 IS A RECEPTOR FOR FIBRONECTIN
CC AND CYTOTACTIN. IT RECOGNIZES THE SEQUENCE R-G-D IT ITS LIGANDS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-6
CC ASSOCIATES WITH ALPHA-V.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
CC -!- SIMILARITY: CONTAINS 1 VWFA-LIKE DOMAIN.
CC ----------

CC ----------
DR EMBL; M35197; AAA37043.1; -.
DR EMBL; A26611; CAA01833.1; -.
DR PIR; B37057; B37057.
DR HSSP; P04355; 2MRT.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002369; Integrin_B.
DR InterPro; IPR001169; Integrin_beta_C.
DR Pfam; PF00362; integrin_B; 1.
DR ProDom; PD001811; Integrin_B; 1.
DR SMART; SM00001; EGF_like; 1.
DR SMART; SM00187; INB; 1.
DR PROSITE; PS00243; INTEGRIN_BETA; 2.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
KW Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;
KW Repeat.
FT NON_TER 1 1
FT DOMAIN <1 566 EXTRACELLULAR (POTENTIAL).

```
FT   TRANSMEM     567   >577       POTENTIAL.
FT   DOMAIN        <1    230       VWFA-LIKE.
FT   DOMAIN       315    478       4 CYSTEINE-RICH TANDEM REPEATS.
FT   REPEAT       315    360       I.
FT   REPEAT       361    402       II.
FT   REPEAT       403    441       III.
FT   REPEAT       442    478       IV.
FT   DISULFID      56     63       BY SIMILARITY.
FT   DISULFID     111    152       BY SIMILARITY.
FT   DISULFID     253    265       BY SIMILARITY.
FT   DISULFID     285    529       BY SIMILARITY.
FT   DISULFID     311    315       BY SIMILARITY.
FT   DISULFID     326    338       BY SIMILARITY.
FT   DISULFID     335    370       BY SIMILARITY.
FT   DISULFID     340    349       BY SIMILARITY.
FT   DISULFID     351    361       BY SIMILARITY.
FT   DISULFID     376    381       BY SIMILARITY.
FT   DISULFID     378    411       BY SIMILARITY.
FT   DISULFID     383    396       BY SIMILARITY.
FT   DISULFID     398    403       BY SIMILARITY.
FT   DISULFID     417    422       BY SIMILARITY.
FT   DISULFID     419    450       BY SIMILARITY.
FT   DISULFID     424    433       BY SIMILARITY.
FT   DISULFID     435    442       BY SIMILARITY.
FT   DISULFID     456    461       BY SIMILARITY.
FT   DISULFID     458    504       BY SIMILARITY.
FT   DISULFID     463    473       BY SIMILARITY.
FT   DISULFID     476    479       BY SIMILARITY.
FT   DISULFID     483    492       BY SIMILARITY.
FT   DISULFID     489    561       BY SIMILARITY.
FT   DISULFID     508    537       BY SIMILARITY.
FT   CARBOHYD     119    119       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     246    246       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     255    255       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     277    277       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     322    322       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     330    330       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     400    400       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     434    434       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     555    555       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   NON_TER      577    577
SQ   SEQUENCE   577 AA;  62298 MW;  B83B468C83EDCCF9 CRC64;


  Query Match             67.7%;  Score 44;  DB 1;  Length 577;
  Best Local Similarity   66.7%;  Pred. No. 7.2;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           |  ||||:|
Db     370 CSGRGDCYC 378


RESULT   2
ITB6_MOUSE
ID   ITB6_MOUSE     STANDARD;      PRT;   787 AA.
AC   Q9Z0T9;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Integrin beta-6 precursor.
GN   ITGB6.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=20547423; PubMed=11095652;
RA   Arend L.J., Smart A.M., Briggs J.P.;
RT   "Mouse beta(6) integrin sequence, pattern of expression, and role in
RT   kidney development.";
RL   J. Am. Soc. Nephrol. 11:2297-2305(2000).
CC   -!- FUNCTION: INTEGRIN ALPHA-V/BETA-6 IS A RECEPTOR FOR FIBRONECTIN
CC       AND CYTOTACTIN. IT RECOGNIZES THE SEQUENCE R-G-D IT ITS LIGANDS
CC       (BY SIMILARITY).
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-6
CC       ASSOCIATES WITH ALPHA-V (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 VWFA-LIKE DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF115376; AAD17212.1; -.
DR   MGD; MGI:96615; Itgb6.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR002369; Integrin_B.
DR   InterPro; IPR001169; Integrin_beta_C.
DR   InterPro; IPR003659; PSI.
DR   InterPro; IPR002035; vWFA.
DR   Pfam; PF00362; integrin_B; 1.
DR   PRINTS; PR01186; INTEGRINB.
DR   ProDom; PD001811; Integrin_B; 1.
DR   SMART; SM00001; EGF_like; 1.
DR   SMART; SM00187; INB; 1.
DR   SMART; SM00423; PSI; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00243; INTEGRIN_BETA; 2.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
KW   Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;
KW   Repeat; Signal.
FT   SIGNAL         1     21       POTENTIAL.
FT   CHAIN         22    787       INTEGRIN BETA-6.
FT   DOMAIN        22    706       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     707    729       POTENTIAL.
FT   DOMAIN       730    787       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       131    371       VWFA-LIKE.
FT   DOMAIN       456    619       4 CYSTEINE-RICH TANDEM REPEATS.
FT   REPEAT       456    501       I.
FT   REPEAT       502    543       II.
FT   REPEAT       544    582       III.
FT   REPEAT       583    619       IV.
FT   DISULFID      23    454       BY SIMILARITY.
FT   DISULFID      31     41       BY SIMILARITY.
FT   DISULFID      34     70       BY SIMILARITY.
FT   DISULFID      44     59       BY SIMILARITY.
FT   DISULFID     197    204       BY SIMILARITY.
FT   DISULFID     252    293       BY SIMILARITY.
FT   DISULFID     394    406       BY SIMILARITY.
FT   DISULFID     426    669       BY SIMILARITY.
FT   DISULFID     452    456       BY SIMILARITY.
FT   DISULFID     467    479       BY SIMILARITY.
FT   DISULFID     476    511       BY SIMILARITY.
FT   DISULFID     481    490       BY SIMILARITY.
FT   DISULFID     492    502       BY SIMILARITY.
FT   DISULFID     517    522       BY SIMILARITY.
FT   DISULFID     519    552       BY SIMILARITY.
FT   DISULFID     524    537       BY SIMILARITY.
FT   DISULFID     539    544       BY SIMILARITY.
FT   DISULFID     558    563       BY SIMILARITY.
FT   DISULFID     560    591       BY SIMILARITY.
FT   DISULFID     565    574       BY SIMILARITY.
FT   DISULFID     576    583       BY SIMILARITY.
FT   DISULFID     597    602       BY SIMILARITY.
FT   DISULFID     599    645       BY SIMILARITY.
FT   DISULFID     604    614       BY SIMILARITY.
FT   DISULFID     617    620       BY SIMILARITY.
```

```
FT   DISULFID    624    633       BY SIMILARITY.
FT   DISULFID    630    701       BY SIMILARITY.
FT   DISULFID    649    677       BY SIMILARITY.
FT   CARBOHYD     48     48       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     97     97       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    260    260       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    387    387       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    418    418       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    463    463       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    471    471       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    541    541       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    575    575       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   787 AA;  86041 MW;  C6438C6F1E6B7FBD CRC64;


  Query Match             67.7%;  Score 44;  DB 1;  Length 787;
  Best Local Similarity   66.7%;  Pred. No. 9.3;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           |  ||||:|
Db     511 CSGRGDCYC 519


RESULT   3
ITB6_HUMAN
ID   ITB6_HUMAN     STANDARD;      PRT;   788 AA.
AC   P18564; Q16500;
DT   01-NOV-1990 (Rel. 16, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Integrin beta-6 precursor.
GN   ITGB6.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pancreas;
RX   MEDLINE=90307659; PubMed=2365683;
RA   Sheppard D., Rozzo C., Starr L., Quaranta V., Erle D.J., Pytela R.;
RT   "Complete amino acid sequence of a novel integrin beta subunit (beta
RT   6) identified in epithelial cells using the polymerase chain
RT   reaction.";
RL   J. Biol. Chem. 265:11502-11507(1990).
RN   [2]
RP   REVISIONS TO 18-24; 158; 642 AND 719.
RA   Askins J.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 116-197 FROM N.A.
RX   MEDLINE=93002753; PubMed=1382574;
RA   Jiang W.M., Jenkins D., Yuan Q., Leung E., Choo K.H., Watson J.D.,
RA   Krissansen G.W.;
RT   "The gene organization of the human beta 7 subunit, the common beta
RT   subunit of the leukocyte integrins HML-1 and LPAM-1.";
RL   Int. Immunol. 4:1031-1040(1992).
CC   -!- FUNCTION: INTEGRIN ALPHA-V/BETA-6 IS A RECEPTOR FOR FIBRONECTIN
CC       AND CYTOTACTIN. IT RECOGNIZES THE SEQUENCE R-G-D IT ITS LIGANDS.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-6
CC       ASSOCIATES WITH ALPHA-V.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 VWFA-LIKE DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M35198; AAA36122.2; -.
DR   EMBL; A26609; CAA01832.1; -.
DR   EMBL; S49380; AAB23690.1; -.
DR   PIR; A37057; A37057.
DR   HSSP; P04355; 2MRT.
DR   MIM; 147558; -.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR002369; Integrin_B.
DR   InterPro; IPR003659; PSI.
DR   InterPro; IPR002035; vWFA.
DR   Pfam; PF00362; integrin_B; 1.
DR   PRINTS; PR01186; INTEGRINB.
DR   ProDom; PD001811; Integrin_B; 1.
DR   SMART; SM00001; EGF_like; 1.
DR   SMART; SM00187; INB; 1.
DR   SMART; SM00423; PSI; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00243; INTEGRIN_BETA; 3.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
KW   Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;
KW   Repeat; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    788       INTEGRIN BETA-6.
FT   DOMAIN       22    707       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    708    730       POTENTIAL.
FT   DOMAIN      731    788       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      131    371       VWFA-LIKE.
FT   DOMAIN      456    619       4 CYSTEINE-RICH TANDEM REPEATS.
FT   REPEAT      456    501       I.
FT   REPEAT      502    543       II.
FT   REPEAT      544    582       III.
FT   REPEAT      583    619       IV.
FT   DISULFID     23    454       BY SIMILARITY.
FT   DISULFID     31     41       BY SIMILARITY.
FT   DISULFID     34     70       BY SIMILARITY.
FT   DISULFID     44     59       BY SIMILARITY.
FT   DISULFID    197    204       BY SIMILARITY.
FT   DISULFID    252    293       BY SIMILARITY.
FT   DISULFID    394    406       BY SIMILARITY.
FT   DISULFID    426    670       BY SIMILARITY.
FT   DISULFID    452    456       BY SIMILARITY.
FT   DISULFID    467    479       BY SIMILARITY.
FT   DISULFID    476    511       BY SIMILARITY.
FT   DISULFID    481    490       BY SIMILARITY.
FT   DISULFID    492    502       BY SIMILARITY.
FT   DISULFID    517    522       BY SIMILARITY.
FT   DISULFID    519    552       BY SIMILARITY.
FT   DISULFID    524    537       BY SIMILARITY.
FT   DISULFID    539    544       BY SIMILARITY.
FT   DISULFID    558    563       BY SIMILARITY.
FT   DISULFID    560    591       BY SIMILARITY.
FT   DISULFID    565    574       BY SIMILARITY.
FT   DISULFID    576    583       BY SIMILARITY.
FT   DISULFID    597    602       BY SIMILARITY.
FT   DISULFID    599    645       BY SIMILARITY.
FT   DISULFID    604    614       BY SIMILARITY.
FT   DISULFID    617    620       BY SIMILARITY.
FT   DISULFID    624    633       BY SIMILARITY.
FT   DISULFID    630    702       BY SIMILARITY.
FT   DISULFID    649    678       BY SIMILARITY.
FT   CARBOHYD     48     48       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     97     97       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    260    260       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    387    387       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    396    396       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    463    463       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    471    471       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    541    541       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    575    575       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   788 AA;  85935 MW;  EDB7D533EC4C8C4D CRC64;
```

```
  Query Match             67.7%;  Score 44;  DB 1;  Length 788;
  Best Local Similarity   66.7%;  Pred. No. 9.3;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           |  ||||:|
Db     511 CSGRGDCYC 519
```

```
RESULT   4
PAPA_HUMAN
ID   PAPA_HUMAN     STANDARD;      PRT;  1627 AA.
AC   Q13219; Q08371; Q9UDK7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Pregnancy-associated plasma protein-A precursor (EC 3.4.24.-) (PAPP-A)
DE   (Insulin-like growth factor-dependent IGF binding protein-4 protease)
DE   (IGF-dependent IGFBP-4 protease) (IGFBP-4ase).
GN   PAPPA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND INDUCTION.
RC   TISSUE=Placenta;
RX   MEDLINE=96203921; PubMed=8620868;
RA   Haaning J., Oxvig C., Overgaard M.T., Ebbesen P., Kristensen T.,
RA   Sottrup-Jensen L.;
RT   "Complete cDNA sequence of the preproform of human pregnancy-
RT   associated plasma protein-A. Evidence for expression in the brain and
RT   induction by cAMP.";
RL   Eur. J. Biochem. 237:159-163(1996).
RN   [2]
RP   SEQUENCE OF 77-1627 FROM N.A., SEQUENCE OF 81-98; 117-126; 210-224;
RP   466-485; 507-519; 576-593; 609-621; 718-736; 742-754; 1006-1017;
RP   1259-1273; 1369-1374; 1389-1398; 1490-1509; 1524-1533 AND 1537-1544,
RP   VARIANT SER-944, AND TISSUE SPECIFICITY.
RC   TISSUE=Placenta, and Serum;
RX   MEDLINE=94146014; PubMed=7508748;
RA   Kristensen T., Oxvig C., Sand O., Moller N.P.H., Sottrup-Jensen L.;
RT   "Amino acid sequence of human pregnancy-associated plasma protein-A
RT   derived from cloned cDNA.";
RL   Biochemistry 33:1592-1598(1994).
RN   [3]
RP   SEQUENCE OF 81-89; 117-126; 210-224; 460-485; 507-519; 576-593;
RP   718-736; 742-754; 1259-1273; 1369-1374; 1490-1509; 1524-1533 AND
RP   1537-1544, SUBUNITS, AND INTERCHAIN DISULFIDE BOND.
RC   TISSUE=Serum;
RX   MEDLINE=93286045; PubMed=7685339;
RA   Oxvig C., Sand O., Kristensen T., Gleich G.J., Sottrup-Jensen L.;
RT   "Circulating human pregnancy-associated plasma protein-A is disulfide-
RT   bridged to the proform of eosinophil major basic protein.";
RL   J. Biol. Chem. 268:12243-12246(1993).
RN   [4]
RP   IDENTIFICATION, FUNCTION, SUBCELLULAR LOCATION, AND TISSUE
RP   SPECIFICITY.
RC   TISSUE=Fibroblast;
RX   MEDLINE=99179030; PubMed=10077652;
RA   Lawrence J.B., Oxvig C., Overgaard M.T., Sottrup-Jensen L.,
RA   Gleich G.J., Hays L.G., Yates J.R. III, Conover C.A.;
RT   "The insulin-like growth factor (IGF)-dependent IGF binding protein-4
RT   protease secreted by human fibroblasts is pregnancy-associated plasma
RT   protein-A.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:3149-3153(1999).
RN   [5]
RP   FUNCTION, SUBUNITS, AND ENZYME REGULATION.
RX   MEDLINE=20469470; PubMed=10913121;
RA   Overgaard M.T., Haaning J., Boldt H.B., Olsen I.M., Laursen L.S.,
RA   Christiansen M., Gleich G.J., Sottrup-Jensen L., Conover C.A.,
RA   Oxvig C.;
RT   "Expression of recombinant human pregnancy-associated plasma protein-A
RT   and identification of the proform of eosinophil major basic protein
RT   as its physiological inhibitor.";
RL   J. Biol. Chem. 275:31128-31133(2000).
RN   [6]
RP   TISSUE SPECIFICITY.
RX   MEDLINE=95057018; PubMed=7526035;
RA   Bonno M., Oxvig C., Kephart G.M., Wagner J.M., Kristensen T.,
RA   Sottrup-Jensen L., Gleich G.J.;
RT   "Localization of pregnancy-associated plasma protein-A and
RT   colocalization of pregnancy-associated plasma protein-A messenger
RT   ribonucleic acid and eosinophil granule major basic protein messenger
RT   ribonucleic acid in placenta.";
RL   Lab. Invest. 71:560-566(1994).
RN   [7]
RP   TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RX   MEDLINE=99423540; PubMed=10491647;
RA   Overgaard M.T., Oxvig C., Christiansen M., Lawrence J.B.,
RA   Conover C.A., Gleich G.J., Sottrup-Jensen L., Haaning J.;
RT   "Messenger ribonucleic acid levels of pregnancy-associated plasma
RT   protein-A and the proform of eosinophil major basic protein:
RT   expression in human reproductive and nonreproductive tissues.";
RL   Biol. Reprod. 61:1083-1089(1999).
RN   [8]
RP   DEVELOPMENTAL STAGE.
RX   MEDLINE=95293954; PubMed=7539791;
RA   Oxvig C., Haaning J., Kristensen L., Wagner J.M., Rubin I.,
RA   Stigbrand T., Gleich G.J., Sottrup-Jensen L.;
RT   "Identification of angiotensinogen and complement C3dg as novel
RT   proteins binding the proform of eosinophil major basic protein in
RT   human pregnancy serum and plasma.";
RL   J. Biol. Chem. 270:13645-13651(1995).
CC   -!- FUNCTION: Metalloproteinase which specifically cleaves IGFBP-4 in
CC       the presence of IGF, resulting in release of bound IGF.
CC   -!- ENZYME REGULATION: Inhibited by complexation with the proform
CC       of PRG2.
CC   -!- SUBUNIT: Homodimer; disulfide-linked. In pregnancy serum,
CC       predominantly found as a disulfide-linked 2:2 heterotetramer with
CC       the proform of PRG2.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: High levels in placenta and pregnancy serum.
CC       In placenta, expressed in X cells in septa and anchoring villi,
CC       and in syncytiotrophoblasts in the chorionic villi. Lower levels
CC       are found in a variety of other tissues including kidney,
CC       myometrium, endometrium, ovaries, breast, prostate, bone marrow,
CC       colon, fibroblasts and osteoblasts.
CC   -!- DEVELOPMENTAL STAGE: Present in serum and placenta during
CC       pregnancy; levels increase throughout pregnancy.
CC   -!- INDUCTION: By 8-bromoadenosine-3',5'-phosphate.
CC   -!- PTM: There appear to be no free cysteinyl groups.
CC   -!- SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M46.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U28727; AAC50543.1; -.
DR   EMBL; X68280; CAA48341.1; -.
DR   MIM; 176385; -.
DR   MEROPS; M46.001; -.
DR   InterPro; IPR000800; Notch.
DR   InterPro; IPR000436; Sushi_SCR_CCP.
DR   InterPro; IPR000130; Zn_MTpeptdse.
DR   Pfam; PF00084; sushi; 4.
DR   SMART; SM00032; CCP; 4.
DR   SMART; SM00004; NL; 2.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
KW   Hydrolase; Metalloprotease; Metal-binding; Zinc; Signal; Glycoprotein;
```

```
KW   Zymogen; Repeat; Sushi.
FT   SIGNAL       1     22       POTENTIAL.
FT   PROPEP      23     80
FT   CHAIN       81   1627       PREGNANCY-ASSOCIATED PLASMA PROTEIN-A.
FT   DOMAIN      24     83       ARG-RICH.
FT   DOMAIN    1215   1280       SUSHI 1.
FT   DOMAIN    1285   1342       SUSHI 2.
FT   DOMAIN    1346   1410       SUSHI 3.
FT   DOMAIN    1415   1471       SUSHI 4.
FT   METAL      562    562       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE   563    563       BY SIMILARITY.
FT   METAL      566    566       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DISULFID   461    461       INTERCHAIN (WITH THE PROFORM OF
FT                               PRG2).
FT   CARBOHYD   390    390       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   402    402       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   429    429       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   480    480       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   601    601       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   619    619       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   725    725       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   825    825       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1026   1026       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1222   1222       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1226   1226       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1323   1323       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1465   1465       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1519   1519       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT    944    944       R -> S.
FT                               /FTId=VAR_011419.
FT   CONFLICT   511    512       TH -> RD (IN REF. 2; AA SEQUENCE).
SQ   SEQUENCE   1627 AA;  181005 MW;  34164AA3EE56C768 CRC64;


  Query Match             66.2%;  Score 43;  DB 1;  Length 1627;
  Best Local Similarity   66.7%;  Pred. No. 24;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 CDCRGDCFC 9
          || :||| |
Db   1600 CDLQGDCAC 1608


RESULT   5
LRP_CAEEL
ID   LRP_CAEEL      STANDARD;      PRT;  4753 AA.
AC   Q04833;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Low-density lipoprotein receptor-related protein precursor (LRP).
GN   LRP OR F29D11.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93281621; PubMed=8506301;
RA   Yochem J., Greenwald I.;
RT   "A gene for a low density lipoprotein receptor-related protein in the
RT   nematode Caenorhabditis elegans.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:4572-4576(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Wilkinson J.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: MAY ACT AS A RECEPTOR FOR THE ENDOCYTOSIS OF
CC       EXTRACELLULAR LIGANDS SUCH AS CHYLOMICRON REMNANTS, PROTEASE-
CC       INHIBITOR COMPLEXES AND VITELLOGENIN.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: CONTAINS 35 LDL-RECEPTOR CLASS A DOMAINS.
CC   -!- SIMILARITY: CONTAINS 17 EGF-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M96150; AAA28105.1; -.
DR   EMBL; Z73907; CAA98124.1; -.
DR   HSSP; Q07954; 1CR8.
DR   WormPep; F29D11.1; CE05765.
DR   InterPro; IPR000152; Asx_hydroxyl.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR002172; LDL_recept_A.
DR   InterPro; IPR000033; Ldl_receptor_rep.
DR   Pfam; PF00057; ldl_recept_a; 35.
DR   Pfam; PF00058; ldl_recept_b; 26.
DR   PRINTS; PR00261; LDLRECEPTOR.
DR   SMART; SM00179; EGF_CA; 2.
DR   SMART; SM00001; EGF_like; 15.
DR   SMART; SM00192; LDLa; 34.
DR   SMART; SM00135; LY; 30.
DR   PROSITE; PS00010; ASX_HYDROXYL; 6.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 3.
DR   PROSITE; PS01187; EGF_CA; 3.
DR   PROSITE; PS01209; LDLRA_1; 27.
DR   PROSITE; PS50068; LDLRA_2; 34.
KW   Receptor; Transmembrane; Repeat; Endocytosis; Glycoprotein;
KW   Signal; Calcium-binding; EGF-like domain; Coated pits.
FT   SIGNAL       1     18       POTENTIAL.
FT   CHAIN       19   4753       LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED
FT                               PROTEIN.
FT   DOMAIN      19   4570       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM  4571   4596       POTENTIAL.
FT   DOMAIN    4597   4753       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      51     89       LDL-RECEPTOR CLASS A 1.
FT   DOMAIN      90    133       LDL-RECEPTOR CLASS A 2.
FT   DOMAIN     136    177       LDL-RECEPTOR CLASS A 3.
FT   DOMAIN     180    220       LDL-RECEPTOR CLASS A 4.
FT   DOMAIN     221    259       LDL-RECEPTOR CLASS A 5.
FT   DOMAIN     260    298       LDL-RECEPTOR CLASS A 6.
FT   DOMAIN     299    337       EGF-LIKE 1.
FT   DOMAIN     338    368       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     669    712       EGF-LIKE 3.
FT   DOMAIN     997   1043       EGF-LIKE 4.
FT   DOMAIN    1052   1097       LDL-RECEPTOR CLASS A 7.
FT   DOMAIN    1099   1140       LDL-RECEPTOR CLASS A 8.
FT   DOMAIN    1144   1184       LDL-RECEPTOR CLASS A 9.
FT   DOMAIN    1185   1225       LDL-RECEPTOR CLASS A 10.
FT   DOMAIN    1226   1265       LDL-RECEPTOR CLASS A 11.
FT   DOMAIN    1268   1309       LDL-RECEPTOR CLASS A 12.
FT   DOMAIN    1311   1352       LDL-RECEPTOR CLASS A 13.
FT   DOMAIN    1357   1397       LDL-RECEPTOR CLASS A 14.
FT   DOMAIN    1398   1436       EGF-LIKE 5.
FT   DOMAIN    1437   1476       EGF-LIKE 6.
FT   DOMAIN    1747   1786       EGF-LIKE 7.
FT   DOMAIN    2080   2120       EGF-LIKE 8.
FT   DOMAIN    2396   2439       EGF-LIKE 9.
FT   DOMAIN    2728   2780       EGF-LIKE 10.
FT   DOMAIN    2790   2831       LDL-RECEPTOR CLASS A 15.
FT   DOMAIN    2832   2870       LDL-RECEPTOR CLASS A 16.
FT   DOMAIN    2872   2914       LDL-RECEPTOR CLASS A 17.
FT   DOMAIN    2917   2958       LDL-RECEPTOR CLASS A 18.
FT   DOMAIN    2959   2999       LDL-RECEPTOR CLASS A 19.
FT   DOMAIN    3004   3046       LDL-RECEPTOR CLASS A 20.
FT   DOMAIN    3047   3095       LDL-RECEPTOR CLASS A 21.
FT   DOMAIN    3098   3137       LDL-RECEPTOR CLASS A 22.
FT   DOMAIN    3138   3176       LDL-RECEPTOR CLASS A 23.
```

```
FT   DOMAIN      3185    3223       LDL-RECEPTOR CLASS A 24.
FT   DOMAIN      3224    3265       EGF-LIKE 11.
FT   DOMAIN      3266    3306       EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      3582    3624       EGF-LIKE 13.
FT   DOMAIN      3625    3668       LDL-RECEPTOR CLASS A 25.
FT   DOMAIN      3669    3707       LDL-RECEPTOR CLASS A 26.
FT   DOMAIN      3707    3748       LDL-RECEPTOR CLASS A 27.
FT   DOMAIN      3751    3790       LDL-RECEPTOR CLASS A 28.
FT   DOMAIN      3791    3832       LDL-RECEPTOR CLASS A 29.
FT   DOMAIN      3831    3873       LDL-RECEPTOR CLASS A 30.
FT   DOMAIN      3876    3914       LDL-RECEPTOR CLASS A 31.
FT   DOMAIN      3915    3953       LDL-RECEPTOR CLASS A 32.
FT   DOMAIN      3957    3997       LDL-RECEPTOR CLASS A 33.
FT   DOMAIN      3998    4042       LDL-RECEPTOR CLASS A 34.
FT   DOMAIN      4047    4085       LDL-RECEPTOR CLASS A 35.
FT   DOMAIN      4088    4131       EGF-LIKE 14.
FT   DOMAIN      4132    4176       EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      4477    4515       EGF-LIKE 16.
FT   DOMAIN      4526    4554       EGF-LIKE 17.
FT   SITE        4653    4658       ENDOCYTOSIS SIGNAL (POTENTIAL).
FT   SITE        4744    4744       CRITICAL FOR ENDOCYTOSIS (BY SIMILARITY).
FT   DISULFID      53      65       BY SIMILARITY.
FT   DISULFID      60      78       BY SIMILARITY.
FT   DISULFID      72      87       BY SIMILARITY.
FT   DISULFID      92     109       BY SIMILARITY.
FT   DISULFID      99     122       BY SIMILARITY.
FT   DISULFID     116     131       BY SIMILARITY.
FT   DISULFID     138     152       BY SIMILARITY.
FT   DISULFID     146     165       BY SIMILARITY.
FT   DISULFID     159     175       BY SIMILARITY.
FT   DISULFID     182     195       BY SIMILARITY.
FT   DISULFID     189     208       BY SIMILARITY.
FT   DISULFID     202     218       BY SIMILARITY.
FT   DISULFID     223     235       BY SIMILARITY.
FT   DISULFID     230     248       BY SIMILARITY.
FT   DISULFID     242     257       BY SIMILARITY.
FT   DISULFID     262     275       BY SIMILARITY.
FT   DISULFID     269     288       BY SIMILARITY.
FT   DISULFID     282     297       BY SIMILARITY.
FT   DISULFID     302     311       BY SIMILARITY.
FT   DISULFID     307     320       BY SIMILARITY.
FT   DISULFID     322     336       BY SIMILARITY.
FT   DISULFID     342     352       BY SIMILARITY.
FT   DISULFID     348     361       BY SIMILARITY.
FT   DISULFID     363     367       BY SIMILARITY.
FT   DISULFID     673     682       BY SIMILARITY.
FT   DISULFID     678     697       BY SIMILARITY.
FT   DISULFID     699     711       BY SIMILARITY.
FT   DISULFID    1001    1010       BY SIMILARITY.
FT   DISULFID    1006    1026       BY SIMILARITY.
FT   DISULFID    1028    1042       BY SIMILARITY.
FT   DISULFID    1054    1068       BY SIMILARITY.
FT   DISULFID    1063    1081       BY SIMILARITY.
FT   DISULFID    1075    1095       BY SIMILARITY.
FT   DISULFID    1101    1114       BY SIMILARITY.
FT   DISULFID    1108    1127       BY SIMILARITY.
FT   DISULFID    1121    1138       BY SIMILARITY.
FT   DISULFID    1146    1158       BY SIMILARITY.
FT   DISULFID    1153    1171       BY SIMILARITY.
FT   DISULFID    1165    1182       BY SIMILARITY.
FT   DISULFID    1187    1199       BY SIMILARITY.
FT   DISULFID    1194    1212       BY SIMILARITY.
FT   DISULFID    1206    1223       BY SIMILARITY.
FT   DISULFID    1228    1241       BY SIMILARITY.
FT   DISULFID    1235    1254       BY SIMILARITY.
FT   DISULFID    1248    1263       BY SIMILARITY.
FT   DISULFID    1270    1283       BY SIMILARITY.
FT   DISULFID    1277    1296       BY SIMILARITY.
FT   DISULFID    1290    1307       BY SIMILARITY.
FT   DISULFID    1313    1325       BY SIMILARITY.
FT   DISULFID    1320    1338       BY SIMILARITY.
FT   DISULFID    1332    1350       BY SIMILARITY.
FT   DISULFID    1359    1373       BY SIMILARITY.
FT   DISULFID    1366    1386       BY SIMILARITY.
FT   DISULFID    1380    1396       BY SIMILARITY.
FT   DISULFID    1401    1412       BY SIMILARITY.
FT   DISULFID    1408    1421       BY SIMILARITY.
FT   DISULFID    1423    1435       BY SIMILARITY.
FT   DISULFID    1441    1451       BY SIMILARITY.
FT   DISULFID    1447    1460       BY SIMILARITY.
FT   DISULFID    1462    1475       BY SIMILARITY.
FT   DISULFID    1751    1760       BY SIMILARITY.
FT   DISULFID    1756    1770       BY SIMILARITY.
FT   DISULFID    1772    1785       BY SIMILARITY.
FT   DISULFID    2084    2095       BY SIMILARITY.
FT   DISULFID    2091    2105       BY SIMILARITY.
FT   DISULFID    2107    2119       BY SIMILARITY.
FT   DISULFID    2400    2415       BY SIMILARITY.
FT   DISULFID    2411    2426       BY SIMILARITY.
FT   DISULFID    2428    2438       BY SIMILARITY.
FT   DISULFID    2732    2743       BY SIMILARITY.
FT   DISULFID    2739    2759       BY SIMILARITY.
FT   DISULFID    2761    2779       BY SIMILARITY.
FT   DISULFID    2792    2805       BY SIMILARITY.
FT   DISULFID    2800    2818       BY SIMILARITY.
FT   DISULFID    2812    2829       BY SIMILARITY.
FT   DISULFID    2834    2846       BY SIMILARITY.
FT   DISULFID    2841    2859       BY SIMILARITY.
FT   DISULFID    2853    2868       BY SIMILARITY.

  Query Match             66.2%;  Score 43;  DB 1;  Length 4753;
  Best Local Similarity   75.0%;  Pred. No. 59;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCF 8
           | | ||||
Db     361 CSCLGDCF 368


RESULT   6
TX25_PHONI
ID   TX25_PHONI     STANDARD;      PRT;    49 AA.
AC   P29424;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Neurotoxin Tx2-5.
OS   Phoneutria nigriventer (Brazilian armed spider).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC   Araneomorphae; Entelegynae; Lycosoidea; Ctenidae; Phoneutria.
OX   NCBI_TaxID=6918;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=93011905; PubMed=1397265;
RA   Cordeiro M.N., Diniz C.R., Valentim A.C., von Eickstedt V.R.D.,
RA   Gilroy J., Richardson M.;
RT   "The purification and amino acid sequences of four Tx2 neurotoxins
RT   from the venom of the Brazilian 'armed' spider Phoneutria nigriventer
RT   (Keys).";
RL   FEBS Lett. 310:153-156(1992).
RN   [2]
RP   SEQUENCE OF 1-10.
RC   TISSUE=Venom;
RX   MEDLINE=92196803; PubMed=1801316;
RA   Rezende L. Jr., Cordeiro M.N., Oliveira E.B., Diniz C.R.;
RT   "Isolation of neurotoxic peptides from the venom of the 'armed'
RT   spider Phoneutria nigriventer.";
RL   Toxicon 29:1225-1233(1991).
CC   -!- FUNCTION: BLOCKS VOLTAGE-GATED SODIUM CHANNELS. CAUSES SCRATCHING,
CC       LACRIMATION, HYPERSALIVATION, SWEATING AND AGITATION FOLLOWED BY
CC       SPASTIC PARALYSIS OF THE ANTERIOR AND POSTERIOR EXTREMITIES AND
CC       DEATH AT DOSE LEVELS OF 0.24 MG/MOUSE. INSECTICIDAL TO THE LARVAL
CC       AND ADULT FORMS OF THE HOUSE FLY.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Produced by the venomous gland.
```

```
CC   -!- SIMILARITY: BELONGS TO THE SPIDER TOXIN TX2 FAMILY.
DR   PIR; B39305; B39305.
DR   PIR; S29215; S29215.
KW   Sodium channel inhibitor; Toxin; Neurotoxin.
SQ   SEQUENCE   49 AA;  5111 MW;  77B46AAAB3911716 CRC64;


  Query Match             65.4%;  Score 42.5;  DB 1;  Length 49;
  Best Local Similarity   58.3%;  Pred. No. 1.5;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    3;  Gaps    1;

Qy      1 CDC---RGDCFC 9
          |||   ||:| |
Db     14 CDCCGERGECVC 25


RESULT   7
TX26_PHONI
ID   TX26_PHONI     STANDARD;      PRT;    82 AA.
AC   P29425; Q95UF2;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Neurotoxin Tx2-6 precursor.
OS   Phoneutria nigriventer (Brazilian armed spider).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC   Araneomorphae; Entelegynae; Lycosoidea; Ctenidae; Phoneutria.
OX   NCBI_TaxID=6918;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RA   Penaforte C., Kalapothakis E.;
RT   "Molecular cloning of Tx2-6, a neurotoxin from the spider Phoneutria
RT   nigriventer.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 35-82 FROM N.A.
RC   TISSUE=Venom;
RX   MEDLINE=93011905; PubMed=1397265;
RA   Cordeiro M.N., Diniz C.R., Valentim A.C., von Eickstedt V.R.D.,
RA   Gilroy J., Richardson M.;
RT   "The purification and amino acid sequences of four Tx2 neurotoxins
RT   from the venom of the Brazilian 'armed' spider Phoneutria nigriventer
RT   (Keys).";
RL   FEBS Lett. 310:153-156(1992).
CC   -!- FUNCTION: BLOCKS VOLTAGE-GATED SODIUM CHANNELS. CAUSES SCRATCHING,
CC       LACRIMATION, HYPERSALIVATION, SWEATING AND AGITATION FOLLOWED BY
CC       SPASTIC PARALYSIS OF THE ANTERIOR AND POSTERIOR EXTREMITIES AND
CC       DEATH AT DOSE LEVELS OF 0.79 MG/MOUSE. IT SIGNIFICANTLY ACTIVATES
CC       VOLTAGE-DEPENDENT SODIUM CHANNELS. INSECTICIDAL TO THE LARVAL AND
CC       ADULT FORMS OF THE HOUSE FLY.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Produced by the venomous gland.
CC   -!- SIMILARITY: BELONGS TO THE SPIDER TOXIN TX2 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY054746; AAL14349.1; -.
DR   PIR; S29216; S29216.
KW   Sodium channel inhibitor; Toxin; Neurotoxin; Signal.
FT   SIGNAL        1     17       POTENTIAL.
FT   PROPEP       18     34
FT   CHAIN        35     81       NEUROTOXIN TX2-6.
FT   PROPEP       82     82
SQ   SEQUENCE   82 AA;  9031 MW;  F4CEA5E7B8D53E59 CRC64;


  Query Match             65.4%;  Score 42.5;  DB 1;  Length 82;
  Best Local Similarity   58.3%;  Pred. No. 2.3;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    3;  Gaps    1;

Qy      1 CDC---RGDCFC 9
          |||   ||:| |
Db     48 CDCCGERGECVC 59


RESULT   8
TX5A_PHONI
ID   TX5A_PHONI     STANDARD;      PRT;    82 AA.
AC   O76199;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Neurotoxin Pn2-5A precursor.
OS   Phoneutria nigriventer (Brazilian armed spider).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC   Araneomorphae; Entelegynae; Lycosoidea; Ctenidae; Phoneutria.
OX   NCBI_TaxID=6918;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RX   MEDLINE=99053403; PubMed=9839668;
RA   Kalapothakis E., Penaforte C.L., Beirao P.S.L., Romano-Silva M.A.,
RA   Cruz J.S., Prado M.A.M., Guimaraes P.E.M., Gomez M.V., Prado V.F.;
RT   "Cloning of cDNAS encoding neurotoxic peptides from the spider
RT   Phoneutria nigriventer.";
RL   Toxicon 36:1843-1850(1998).
CC   -!- FUNCTION: Blocks voltage-gated sodium channels (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Produced by the venomous gland.
CC   -!- SIMILARITY: BELONGS TO THE SPIDER TOXIN TX2 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF014463; AAC26165.1; -.
KW   Sodium channel inhibitor; Toxin; Neurotoxin; Signal.
FT   SIGNAL        1     17       POTENTIAL.
FT   PROPEP       18     34       POTENTIAL.
FT   CHAIN        35     81       NEUROTOXIN PN2-5A.
FT   PROPEP       82     82       POTENTIAL.
SQ   SEQUENCE   82 AA;  8856 MW;  11DAF1EBE78B318F CRC64;


  Query Match             65.4%;  Score 42.5;  DB 1;  Length 82;
  Best Local Similarity   58.3%;  Pred. No. 2.3;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    3;  Gaps    1;

Qy      1 CDC---RGDCFC 9
          |||   ||:| |
Db     48 CDCCGERGECVC 59


RESULT   9
TX21_PHONI
ID   TX21_PHONI     STANDARD;      PRT;    88 AA.
AC   P29423;
DT   01-APR-1993 (Rel. 25, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Neurotoxin Tx2-1 precursor.
OS   Phoneutria nigriventer (Brazilian armed spider).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC   Araneomorphae; Entelegynae; Lycosoidea; Ctenidae; Phoneutria.
OX   NCBI_TaxID=6918;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RX   MEDLINE=99053403; PubMed=9839668;
RA   Kalapothakis E., Penaforte C.L., Beirao P.S.L., Romano-Silva M.A.,
RA   Cruz J.S., Prado M.A.M., Guimaraes P.E.M., Gomez M.V., Prado V.F.;
RT   "Cloning of cDNAS encoding neurotoxic peptides from the spider
RT   Phoneutria nigriventer.";
RL   Toxicon 36:1843-1850(1998).
RN   [2]
RP   SEQUENCE OF 35-87.
RC   TISSUE=Venom;
RX   MEDLINE=93011905; PubMed=1397265;
RA   Cordeiro M.N., Diniz C.R., Valentim A.C., von Eickstedt V.R.D.,
RA   Gilroy J., Richardson M.;
RT   "The purification and amino acid sequences of four Tx2 neurotoxins
RT   from the venom of the Brazilian 'armed' spider Phoneutria nigriventer
RT   (Keys).";
RL   FEBS Lett. 310:153-156(1992).
CC   -!- FUNCTION: BLOCKS VOLTAGE-GATED SODIUM CHANNELS. CAUSES SCRATCHING,
CC       LACRIMATION, HYPERSALIVATION, SWEATING AND AGITATION FOLLOWED BY
CC       SPASTIC PARALYSIS OF THE ANTERIOR AND POSTERIOR EXTREMITIES AND
CC       DEATH AT DOSE LEVELS OF 1.62 MG/MOUSE. INSECTICIDAL TO THE LARVAL
CC       AND ADULT FORMS OF THE HOUSE FLY.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Produced by the venomous gland.
CC   -!- SIMILARITY: BELONGS TO THE SPIDER TOXIN TX2 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF014464; AAC26166.1; -.
DR   PIR; S29214; S29214.
KW   Sodium channel inhibitor; Toxin; Neurotoxin; Signal.
FT   SIGNAL        1     17       POTENTIAL.
FT   PROPEP       18     34
FT   CHAIN        35     87       NEUROTOXIN TX2-1.
FT   PROPEP       88     88
SQ   SEQUENCE   88 AA;  9841 MW;  D8AD07C6A769E647 CRC64;


  Query Match             65.4%;  Score 42.5;  DB 1;  Length 88;
  Best Local Similarity   58.3%;  Pred. No. 2.5;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    3;  Gaps    1;

Qy      1 CDC---RGDCFC 9
          |||   ||:| |
Db     48 CDCCGERGECVC 59


RESULT  10
TX1A_PHONI
ID   TX1A_PHONI     STANDARD;      PRT;   115 AA.
AC   O76198;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Neurotoxin Pn2-1A precursor.
OS   Phoneutria nigriventer (Brazilian armed spider).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC   Araneomorphae; Entelegynae; Lycosoidea; Ctenidae; Phoneutria.
OX   NCBI_TaxID=6918;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RX   MEDLINE=99053403; PubMed=9839668;
RA   Kalapothakis E., Penaforte C.L., Beirao P.S.L., Romano-Silva M.A.,
RA   Cruz J.S., Prado M.A.M., Guimaraes P.E.M., Gomez M.V., Prado V.F.;
RT   "Cloning of cDNAS encoding neurotoxic peptides from the spider
RT   Phoneutria nigriventer.";
RL   Toxicon 36:1843-1850(1998).
CC   -!- FUNCTION: Blocks voltage-gated sodium channels (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Produced by the venomous gland.
CC   -!- SIMILARITY: BELONGS TO THE SPIDER TOXIN TX2 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bicinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF014462; AAC26164.1; -.
KW   Sodium channel inhibitor; Toxin; Neurotoxin; Signal.
FT   SIGNAL        1     17       POTENTIAL.
FT   PROPEP       18     61       POTENTIAL.
FT   CHAIN        62    114       NEUROTOXIN PN2-1A.
FT   PROPEP      115    115       POTENTIAL.
SQ   SEQUENCE   115 AA;  12858 MW;  B7D3321750F7BA50 CRC64;


  Query Match             65.4%;  Score 42.5;  DB 1;  Length 115;
  Best Local Similarity   58.3%;  Pred. No. 3.1;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    3;  Gaps    1;

Qy      1 CDC---RGDCFC 9
          |||   ||:| |
Db     75 CDCCGERGECVC 86


RESULT  11
CRS5_YEAST
ID   CRS5_YEAST     STANDARD;      PRT;    69 AA.
AC   P41902;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Metallothionein-like protein CRS5.
GN   CRS5 OR YOR031W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95014318; PubMed=7929222;
RA   Culotta V.C., Howard W.R., Liu X.F.;
RT   "CRS5 encodes a metallothionein-like protein in Saccharomyces
RT   cerevisiae.";
RL   J. Biol. Chem. 269:25295-25302(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / FY1679;
RA   de Haan M., Maarse A.C., Grivell L.A.;
RL   Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: CRITICAL ROLE IN COPPER (SPECIFIC) HOMEOSTASIS AND
CC       DETOXIFICATION. MAY PROTECT BY DIRECTLY CHELATING AND SEQUESTERING
CC       COPPER IONS.
CC   -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 13.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L29056; AAA66061.1; -.
```

```
DR   EMBL; X87331; -; NOT_ANNOTATED_CDS.
DR   SGD; S0005557; CRS5.
KW   Metal-binding; Metal-thiolate cluster; Chelation.
SQ   SEQUENCE   69 AA;  7321 MW;  CEEF91203A813FF4 CRC64;


  Query Match             63.1%;  Score 41;  DB 1;  Length 69;
  Best Local Similarity   71.4%;  Pred. No. 3.4;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CDCRGDC 7
             ||| |:|
Db       6 CDCEGEC 12


RESULT  12
LML2_CAEEL
ID   LML2_CAEEL     STANDARD;      PRT;  3672 AA.
AC   Q21313;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Laminin-like protein K08C7.3 precursor.
GN   K08C7.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Berks M.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC   -!- SIMILARITY: CONTAINS 21.5 LAMININ EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
CC   -!- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z70286; CAA94293.1; -.
DR   HSSP; P02468; 1KLO.
DR   WormPep; K08C7.3; CE06136.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR001886; LamNT.
DR   InterPro; IPR000034; Laminin_B.
DR   InterPro; IPR002049; Laminin_EGF.
DR   InterPro; IPR001791; Laminin_G.
DR   Pfam; PF00052; laminin_B; 1.
DR   Pfam; PF00053; laminin_EGF; 21.
DR   Pfam; PF00054; laminin_G; 5.
DR   Pfam; PF00055; laminin_Nterm; 1.
DR   PRINTS; PR00011; EGFLAMININ.
DR   ProDom; PD002082; LamNT; 1.
DR   ProDom; PD003031; Laminin_B; 1.
DR   SMART; SM00180; EGF_Lam; 21.
DR   SMART; SM00281; LamB; 1.
DR   SMART; SM00282; LamG; 5.
DR   SMART; SM00136; LamNT; 1.
DR   PROSITE; PS00022; EGF_1; 19.
DR   PROSITE; PS01186; EGF_2; 4.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; 21.
DR   PROSITE; PS50025; LAM_G_DOMAIN; 5.
KW   Hypothetical protein; Laminin EGF-like domain; Signal; Repeat.
FT   SIGNAL        1     27       POTENTIAL.
FT   CHAIN        28   3672       LAMININ-LIKE PROTEIN K08C7.3.
FT   DOMAIN       28    297       LAMININ N-TERMINAL (DOMAIN VI).
FT   DOMAIN      298    356       LAMININ EGF-LIKE 1.
FT   DOMAIN      357    426       LAMININ EGF-LIKE 2.
FT   DOMAIN      427    471       LAMININ EGF-LIKE 3.
FT   DOMAIN      472    518       LAMININ EGF-LIKE 4.
FT   DOMAIN      519    563       LAMININ EGF-LIKE 5.
FT   DOMAIN      564    609       LAMININ EGF-LIKE 6.
FT   DOMAIN      610    655       LAMININ EGF-LIKE 7.
FT   DOMAIN      656    700       LAMININ EGF-LIKE 8.
FT   DOMAIN      701    755       LAMININ EGF-LIKE 9.
FT   DOMAIN      756    808       LAMININ EGF-LIKE 10.
FT   DOMAIN      809    839       LAMININ EGF-LIKE 11 (INCOMPLETE).
FT   DOMAIN     1415   1460       LAMININ EGF-LIKE 12.
FT   DOMAIN     1461   1505       LAMININ EGF-LIKE 13.
FT   DOMAIN     1506   1553       LAMININ EGF-LIKE 14.
FT   DOMAIN     1554   1604       LAMININ EGF-LIKE 15.
FT   DOMAIN     1605   1614       LAMININ EGF-LIKE 16 (N-TERMINAL).
FT   DOMAIN     1615   1796       LAMININ DOMAIN IV.
FT   DOMAIN     1797   1829       LAMININ EGF-LIKE 16 (C-TERMINAL).
FT   DOMAIN     1830   1879       LAMININ EGF-LIKE 17.
FT   DOMAIN     1880   1936       LAMININ EGF-LIKE 18.
FT   DOMAIN     1937   1989       LAMININ EGF-LIKE 19.
FT   DOMAIN     1990   2036       LAMININ EGF-LIKE 20.
FT   DOMAIN     2037   2083       LAMININ EGF-LIKE 21.
FT   DOMAIN     2084   2131       LAMININ EGF-LIKE 22.
FT   DOMAIN     2693   2884       LAMININ G-LIKE 1.
FT   DOMAIN     2896   3066       LAMININ G-LIKE 2.
FT   DOMAIN     3072   3235       LAMININ G-LIKE 3.
FT   DOMAIN     3310   3482       LAMININ G-LIKE 4.
FT   DOMAIN     3488   3669       LAMININ G-LIKE 5.
FT   DISULFID    298    307       BY SIMILARITY.
FT   DISULFID    300    320       BY SIMILARITY.
FT   DISULFID    322    331       BY SIMILARITY.
FT   DISULFID    334    354       BY SIMILARITY.
FT   DISULFID    357    366       BY SIMILARITY.
FT   DISULFID    359    391       BY SIMILARITY.
FT   DISULFID    394    403       BY SIMILARITY.
FT   DISULFID    406    424       BY SIMILARITY.
FT   DISULFID    427    438       BY SIMILARITY.
FT   DISULFID    429    445       BY SIMILARITY.
FT   DISULFID    447    456       BY SIMILARITY.
FT   DISULFID    459    469       BY SIMILARITY.
FT   DISULFID    472    484       BY SIMILARITY.
FT   DISULFID    474    491       BY SIMILARITY.
FT   DISULFID    493    502       BY SIMILARITY.
FT   DISULFID    505    516       BY SIMILARITY.
FT   DISULFID    519    531       BY SIMILARITY.
FT   DISULFID    521    538       BY SIMILARITY.
FT   DISULFID    540    549       BY SIMILARITY.
FT   DISULFID    552    561       BY SIMILARITY.
FT   DISULFID    564    576       BY SIMILARITY.
FT   DISULFID    566    583       BY SIMILARITY.
FT   DISULFID    585    594       BY SIMILARITY.
FT   DISULFID    597    607       BY SIMILARITY.
FT   DISULFID    610    622       BY SIMILARITY.
FT   DISULFID    612    629       BY SIMILARITY.
FT   DISULFID    631    640       BY SIMILARITY.
FT   DISULFID    643    653       BY SIMILARITY.
FT   DISULFID    656    668       BY SIMILARITY.
FT   DISULFID    658    674       BY SIMILARITY.
FT   DISULFID    676    685       BY SIMILARITY.
FT   DISULFID    688    698       BY SIMILARITY.
FT   DISULFID    701    715       BY SIMILARITY.
FT   DISULFID    703    724       BY SIMILARITY.
FT   DISULFID    726    735       BY SIMILARITY.
FT   DISULFID    738    753       BY SIMILARITY.
FT   DISULFID    756    770       BY SIMILARITY.
FT   DISULFID    758    777       BY SIMILARITY.
FT   DISULFID    779    788       BY SIMILARITY.
FT   DISULFID    791    806       BY SIMILARITY.
FT   DISULFID   1415   1427       BY SIMILARITY.
FT   DISULFID   1417   1434       BY SIMILARITY.
FT   DISULFID   1436   1445       BY SIMILARITY.
FT   DISULFID   1448   1458       BY SIMILARITY.
```

```
FT   DISULFID   1461   1469       BY SIMILARITY.
FT   DISULFID   1463   1476       BY SIMILARITY.
FT   DISULFID   1478   1487       BY SIMILARITY.
FT   DISULFID   1490   1503       BY SIMILARITY.
FT   DISULFID   1506   1520       BY SIMILARITY.
FT   DISULFID   1508   1527       BY SIMILARITY.
FT   DISULFID   1529   1538       BY SIMILARITY.
FT   DISULFID   1541   1551       BY SIMILARITY.
FT   DISULFID   1554   1566       BY SIMILARITY.
FT   DISULFID   1556   1573       BY SIMILARITY.
FT   DISULFID   1575   1584       BY SIMILARITY.
FT   DISULFID   1587   1602       BY SIMILARITY.
FT   DISULFID   1830   1839       BY SIMILARITY.
FT   DISULFID   1832   1846       BY SIMILARITY.
FT   DISULFID   1849   1858       BY SIMILARITY.
FT   DISULFID   1861   1877       BY SIMILARITY.
FT   DISULFID   1880   1894       BY SIMILARITY.
FT   DISULFID   1882   1905       BY SIMILARITY.
FT   DISULFID   1907   1916       BY SIMILARITY.
FT   DISULFID   1919   1934       BY SIMILARITY.
FT   DISULFID   1937   1951       BY SIMILARITY.
FT   DISULFID   1939   1958       BY SIMILARITY.
FT   DISULFID   1961   1970       BY SIMILARITY.
FT   DISULFID   1973   1987       BY SIMILARITY.
FT   DISULFID   1990   2000       BY SIMILARITY.
FT   DISULFID   1992   2007       BY SIMILARITY.
FT   DISULFID   2009   2018       BY SIMILARITY.
FT   DISULFID   2021   2034       BY SIMILARITY.
FT   DISULFID   2037   2048       BY SIMILARITY.
FT   DISULFID   2039   2055       BY SIMILARITY.
FT   DISULFID   2057   2066       BY SIMILARITY.
FT   DISULFID   2069   2081       BY SIMILARITY.
FT   DISULFID   2084   2096       BY SIMILARITY.
FT   DISULFID   2086   2103       BY SIMILARITY.
FT   DISULFID   2105   2114       BY SIMILARITY.
FT   DISULFID   2117   2129       BY SIMILARITY.
FT   CARBOHYD    121    121       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    140    140       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    249    249       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    351    351       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    477    477       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    511    511       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    530    530       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    634    634       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    761    761       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1014   1014       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1341   1341       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1705   1705       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1756   1756       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1868   1868       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1944   1944       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1986   1986       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2002   2002       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2159   2159       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2207   2207       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2231   2231       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2235   2235       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2401   2401       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2421   2421       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2487   2487       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2821   2821       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3087   3087       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3242   3242       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   3541   3541       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   3672 AA;  404223 MW;  28E262DB5FF14BFA CRC64;

  Query Match          63.1%;  Score 41;  DB 1;  Length 3672;
  Best Local Similarity  55.6%;  Pred. No. 94;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           ||  | | :|
Db     668 CDSNGQCYC 676


RESULT 13
ITBN_DROME
ID   ITBN_DROME     STANDARD;      PRT;   799 AA.
AC   Q27591; Q9VIG7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Integrin beta-nu precursor.
GN   BETA-INT-NU OR CG1762.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Midgut endoderm;
RX   MEDLINE=94357079; PubMed=8076521;
RA   Yee G.H., Hynes R.O.;
RT   "A novel, tissue-specific integrin subunit, beta nu, expressed in the
RT   midgut of Drosophila melanogaster.";
RL   Development 118:845-858(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- FUNCTION: PROBABLY PLAYS A ROLE IN CELL ADHESION.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 VWFA DOMAIN.
CC   ---------------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; L13305; AAC37169.1; -.
DR   EMBL; AE003669; AAF53952.1; -.
DR   FlyBase; FBgn0010395; beta-Int-nu.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR002369; Integrin_B.
DR   InterPro; IPR002049; Laminin_EGF.
DR   InterPro; IPR003659; PSI.
DR   InterPro; IPR002035; vWFA.
DR   Pfam; PF00362; integrin_B; 1.
DR   PRINTS; PR00011; EGFLAMININ.
DR   ProDom; PD001811; Integrin_B; 1.
DR   SMART; SM00181; EGF; 1.
DR   SMART; SM00187; INB; 1.
DR   SMART; SM00423; PSI; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_4.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_4.
DR   PROSITE; PS00243; INTEGRIN_BETA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;
KW   Signal.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27    799       INTEGRIN BETA-NU.
FT   DOMAIN       27    725       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    726    746       POTENTIAL.
FT   DOMAIN      747    799       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      136    372       VWFA.
FT   CARBOHYD     73     73       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    167    167       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    409    409       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    505    505       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    655    655       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    680    680       E -> G (IN REF. 1).
FT   CONFLICT    701    701       V -> A (IN REF. 1).
SQ   SEQUENCE   799 AA;  90841 MW;  351869D523F07DEB CRC64;


  Query Match             62.3%;  Score 40.5;  DB 1;  Length 799;
  Best Local Similarity   38.9%;  Pred. No. 31;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    9;  Gaps    1;

Qy      1 CDCR---------GDCFC 9
          |:||         ||||
Db    552 CECRECLDCDEKLADCFC 569


RESULT  14
TASM_POVMA
ID   TASM_POVMA     STANDARD;      PRT;   195 AA.
AC   P03078;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Small T antigen.
OS   Mouse polyomavirus (strain A2), and
OS   Mouse polyomavirus (strain 3).
OC   Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX   NCBI_TaxID=10636, 10635;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A2;
RX   MEDLINE=80099647; PubMed=6243401;
RA   Soeda E., Arrand J.R., Smolar N., Walsh J.E., Griffin B.E.;
RT   "Coding potential and regulatory signals of the polyoma virus
RT   genome.";
RL   Nature 283:445-453(1980).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=3;
RX   MEDLINE=80001963; PubMed=225042;
RA   Friedmann T., Esty A., Laporte P., Deininger P.L.;
RT   "The nucleotide sequence and genome organization of the polyoma early
RT   region: extensive nucleotide and amino acid homology with SV40.";
RL   Cell 17:715-724(1979).
CC   -!- SIMILARITY: CONTAINS 1 J DOMAIN.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; J02289; AAA46874.1; -.
DR   EMBL; J02288; AAB59899.1; -.
DR   PIR; A03614; TVVPA.
DR   InterPro; IPR001623; DnaJ_N.
DR   InterPro; IPR003354; Papo_T_antigen.
DR   Pfam; PF00226; DnaJ; 1.
DR   Pfam; PF02380; Papo_T_antigen; 1.
DR   SMART; SM00271; DnaJ; 1.
DR   PROSITE; PS00636; DNAJ_1; FALSE_NEG.
DR   PROSITE; PS50076; DNAJ_2; FALSE_NEG.
KW   Early protein.
FT   DOMAIN       12     75       J-DOMAIN.
SQ   SEQUENCE   195 AA;  22811 MW;  44EED6711E1AFEC3 CRC64;


  Query Match             61.5%;  Score 40;  DB 1;  Length 195;
  Best Local Similarity   53.8%;  Pred. No. 11;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    4;  Gaps    1;

Qy      1 CDCR----GDCFC 9
          || |    |:|||
Db    138 CDARCLVLGECFC 150


RESULT  15
TAMI_POVMA
ID   TAMI_POVMA     STANDARD;      PRT;   421 AA.
AC   P03077;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Middle T antigen.
OS   Mouse polyomavirus (strain A2).
OC   Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX   NCBI_TaxID=10636;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=80099647; PubMed=6243401;
RA   Soeda E., Arrand J.R., Smolar N., Walsh J.E., Griffin B.E.;
RT   "Coding potential and regulatory signals of the polyoma virus
RT   genome.";
RL   Nature 283:445-453(1980).
CC   -!- SIMILARITY: CONTAINS 1 J DOMAIN.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; J02288; AAB59900.1; -.
DR   PIR; A03613; TVVPM.
```

```
DR   InterPro; IPR001623; DnaJ_N.
DR   InterPro; IPR003354; Papo_T_antigen.
DR   Pfam; PF00226; DnaJ; 1.
DR   Pfam; PF02380; Papo_T_antigen; 1.
DR   SMART; SM00271; DnaJ; 1.
DR   PROSITE; PS00636; DNAJ_1; FALSE_NEG.
DR   PROSITE; PS50076; DNAJ_2; FALSE_NEG.
KW   Early protein.
FT   DOMAIN       12     75       J-DOMAIN.
SQ   SEQUENCE   421 AA;  48622 MW;  CA0C25C4984CACB7 CRC64;


  Query Match          61.5%;  Score 40;  DB 1;  Length 421;
  Best Local Similarity   53.8%;  Pred. No. 22;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    4;  Gaps    1;

Qy       1 CDCR----GDCFC 9
           || |    |:|||
Db     138 CDARCLVLGECFC 150


Search completed: May 29, 2002, 09:56:26
Job time: 225 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 29, 2002, 09:52:06 ; Search time 25.98 Seconds
(without alignments)
59.929 Million cell updates/sec

Title: US-09-734-628-1
Perfect score: 65
Sequence: 1 CDCRGDCFC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	44	67.7	114	11	Q9R151	Q9r151 cavia porce
2	43	66.2	920	5	Q9GYG8	Q9gyg8 caenorhabdi
3	42.5	65.4	82	5	O76199	O76199 phoneutria
4	42.5	65.4	82	5	Q95UF2	Q95uf2 phoneutria
5	42.5	65.4	115	5	O76198	O76198 phoneutria
6	42	64.6	463	10	Q9C599	Q9c599 arabidopsis
7	41	63.1	458	17	Q9HPQ4	Q9hpq4 halobacteri
8	41	63.1	535	4	Q96EB1	Q96eb1 homo sapien
9	41	63.1	736	10	Q9SVX7	Q9svx7 arabidopsis
10	41	63.1	3704	5	P91904	P91904 caenorhabdi
11	40	61.5	100	5	Q962G0	Q962g0 littorina l
12	40	61.5	116	12	Q9IBN0	Q9ibn0 polyomaviru
13	40	61.5	119	12	Q84251	Q84251 polyomaviru
14	40	61.5	167	12	Q84326	Q84326 polyomaviru
15	40	61.5	195	11	Q04190	Q04190 mus musculu
16	40	61.5	211	12	Q84854	Q84854 polyomaviru
17	40	61.5	214	12	Q84252	Q84252 polyomaviru
18	40	61.5	313	5	Q24330	Q24330 dictyosteli
19	40	61.5	421	11	Q04188	Q04188 mus musculu
20	40	61.5	421	12	Q89765	Q89765 polyomaviru
21	40	61.5	494	11	Q9QZE6	Q9qze6 mus musculu
22	40	61.5	625	10	Q9LP82	Q9lp82 arabidopsis
23	40	61.5	768	13	Q98TH8	Q98th8 cyprinus ca
24	40	61.5	772	13	Q9PUU4	Q9puu4 ictalurus p
25	40	61.5	806	5	O61677	O61677 lytechinus
26	40	61.5	864	13	Q90Z37	Q90z37 brachydanio
27	39.5	60.8	49	11	Q9WUY0	Q9wuy0 rattus norv
28	39.5	60.8	423	11	Q9ER58	Q9er58 mus musculu
29	39	60.0	30	10	Q9S747	Q9s747 momordica c
30	39	60.0	40	13	Q98TP9	Q98tp9 platichthys
31	39	60.0	48	13	Q98TC0	Q98tc0 seriola qui
32	39	60.0	50	3	Q9UTC0	Q9utc0 schizosacch
33	39	60.0	57	5	Q9N9H2	Q9n9h2 ruditapes p
34	39	60.0	59	5	Q9N9H1	Q9n9h1 ruditapes d
35	39	60.0	65	5	Q95P72	Q95p72 macrobrachi
36	39	60.0	70	5	Q967T9	Q967t9 anadara gra
37	39	60.0	80	5	Q9BIV4	Q9biv4 crassostrea
38	39	60.0	107	5	Q9NG19	Q9ng19 crassostrea
39	39	60.0	311	11	Q9WVG0	Q9wvg0 mus musculu
40	39	60.0	499	11	O88714	O88714 mus musculu
41	39	60.0	505	5	Q25431	Q25431 lytechinus
42	39	60.0	567	12	Q9E7Q7	Q9e7q7 influenza a
43	39	60.0	624	10	Q9FFX9	Q9ffx9 arabidopsis
44	39	60.0	624	10	Q9C5P3	Q9c5p3 arabidopsis
45	39	60.0	677	11	Q923H5	Q923h5 mus musculu

ALIGNMENTS

RESULT 1
Q9R151
ID Q9R151 PRELIMINARY; PRT; 114 AA.
AC Q9R151;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE INTEGRIN BETA 6 (FRAGMENT).
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HARTLEY; TISSUE=TRACHEA;
RA Morishima Y., Uchida Y., Nomura A., Ishii Y., Sakamoto T.,
RA Sekizawa K.;
RT "Guinea-pig beta-6 integrin expression in injured tracheal
RT epithelium.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF169344; AAD49344.1; -.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002369; Integrin_B.
DR Pfam; PF00362; integrin_B; 1.
DR ProDom; PD001811; Integrin_B; 1.
DR SMART; SM00001; EGF_like; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
FT NON_TER 1 1
FT NON_TER 114 114
SQ SEQUENCE 114 AA; 12121 MW; 95D4528EBD0435EF CRC64;

Query Match 67.7%; Score 44; DB 11; Length 114;
Best Local Similarity 66.7%; Pred. No. 1.4;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
| |:|:|
Db 104 CSGRGDCYC 112

```
RESULT   2
Q9GYG8
ID   Q9GYG8      PRELIMINARY;      PRT;   920 AA.
AC   Q9GYG8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 103.9 KDA PROTEIN.
GN   W01C8.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Nhan M.;
RT   "The sequence of C. elegans cosmid W01C8.";
RL   Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 1 SET DOMAIN.
DR   EMBL; U41508; AAG00027.1; -.
DR   InterPro; IPR001214; SET.
DR   Pfam; PF00856; SET; 1.
DR   SMART; SM00317; SET; 1.
DR   PROSITE; PS50280; SET; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   920 AA;  103931 MW;  02999AB2367C4F6A CRC64;

  Query Match           66.2%;  Score 43;  DB 5;  Length 920;
  Best Local Similarity  85.7%;  Pred. No. 13;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CDCRGDC 7
           | |||||
Db     110 CGCRGDC 116


RESULT   3
O76199
ID   O76199      PRELIMINARY;      PRT;    82 AA.
AC   O76199;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   NEUROTOXIN TX2-5A PRECURSOR.
GN   TX2-5A OR PN2-5A.
OS   Phoneutria nigriventer (Brazilian armed spider).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC   Araneomorphae; Entelegynae; Lycosoidea; Ctenidae; Phoneutria.
OX   NCBI_TaxID=6918;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=VENOM GLAND;
RX   MEDLINE=99053403; PubMed=9839668;
RA   Kalapothakis E., Penaforte C.L., Beirao P.S.L., Romano-Silva M.A.,
RA   Cruz J.S., Prado M.A.M., Guimaraes P.E.M., Gomez M.V., Prado V.F.;
RT   "Cloning of cDNAS encoding neurotoxic peptides from the spider
RT   Phoneutria nigriventer.";
RL   Toxicon 36:1843-1850(1998).
CC   -!- SIMILARITY: TO NEUROTOXINS TX2-1 AND TX2-6.
DR   EMBL; AF014463; AAC26165.1; -.
KW   Venom; Neurotoxin; Signal.
FT   SIGNAL        1     17       POTENTIAL.
FT   PROPEP       18     34       POTENTIAL.
FT   CHAIN        35     82       NEUROTOXIN TX2-5A.
SQ   SEQUENCE   82 AA;  8856 MW;  11DAF1EBE78B318F CRC64;

  Query Match           65.4%;  Score 42.5;  DB 5;  Length 82;
  Best Local Similarity  58.3%;  Pred. No. 1.9;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    3;  Gaps    1;

Qy       1 CDC---RGDCFC 9
           |||   ||:| |
Db      48 CDCCGERGECVC 59


RESULT   4
Q95UF2
ID   Q95UF2      PRELIMINARY;      PRT;    82 AA.
AC   Q95UF2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NEUROTOXIN TX2-6.
OS   Phoneutria nigriventer (Brazilian armed spider).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC   Araneomorphae; Entelegynae; Lycosoidea; Ctenidae; Phoneutria.
OX   NCBI_TaxID=6918;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=VENOM GLAND;
RA   Penaforte C., Kalapothakis E.;
RT   "Molecular cloning of Tx2-6, a neurotoxin from the spider Phoneutria
RT   nigriventer.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY054746; AAL14349.1; -.
SQ   SEQUENCE   82 AA;  9031 MW;  F4CEA5E7B8D53E59 CRC64;

  Query Match           65.4%;  Score 42.5;  DB 5;  Length 82;
  Best Local Similarity  58.3%;  Pred. No. 1.9;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    3;  Gaps    1;

Qy       1 CDC---RGDCFC 9
           |||   ||:| |
Db      48 CDCCGERGECVC 59


RESULT   5
O76198
ID   O76198      PRELIMINARY;      PRT;   115 AA.
AC   O76198;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   NEUROTOXIN TX2-1A PRECURSOR.
GN   TX2-1A OR PN2-1A.
OS   Phoneutria nigriventer (Brazilian armed spider).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC   Araneomorphae; Entelegynae; Lycosoidea; Ctenidae; Phoneutria.
OX   NCBI_TaxID=6918;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=VENOM GLAND;
RX   MEDLINE=99053403; PubMed=9839668;
RA   Kalapothakis E., Penaforte C.L., Beirao P.S.L., Romano-Silva M.A.,
RA   Cruz J.S., Prado M.A.M., Guimaraes P.E.M., Gomez M.V., Prado V.F.;
RT   "Cloning of cDNAS encoding neurotoxic peptides from the spider
```

```
RT   Phoneutria nigriventer.";
RL   Toxicon 36:1843-1850(1998).
CC   -!- SIMILARITY: TO NEUROTOXINS TX2-5 AND TX2-6.
DR   EMBL; AF014462; AAC26164.1; -.
KW   Venom; Neurotoxin; Signal.
FT   SIGNAL        1     17       POTENTIAL.
FT   PROPEP       18     61       POTENTIAL.
FT   CHAIN        62    114       NEUROTOXIN TX2-1A.
FT   PROPEP      115    115       POTENTIAL.
SQ   SEQUENCE   115 AA;  12858 MW;  B7D3321750F7BA50 CRC64;
```

```
  Query Match             65.4%;  Score 42.5;  DB 5;  Length 115;
  Best Local Similarity   58.3%;  Pred. No. 2.6;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    3;  Gaps    1;

Qy      1 CDC---RGDCFC 9
          |||   ||:| |
Db     75 CDCCGERGECVC 86
```

```
RESULT   6
Q9C599
ID   Q9C599      PRELIMINARY;      PRT;   463 AA.
AC   Q9C599;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 52.0 KDA PROTEIN.
GN   AT5G08780.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
RA   Rudd S., Lemcke K., Mayer K.F.X.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL590346; CAC35883.1; -.
DR   HSSP; P02259; 1HST.
DR   InterPro; IPR001386; Linker_histone.
DR   Pfam; PF00538; linker_histone; 1.
DR   SMART; SM00526; H15; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   463 AA;  52015 MW;  781A08F0BB11DCAA CRC64;
```

```
  Query Match             64.6%;  Score 42;  DB 10;  Length 463;
  Best Local Similarity   62.5%;  Pred. No. 11;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 CDCRGDCF 8
          |||  ||:
Db    114 CDCNNDCY 121
```

```
RESULT   7
Q9HPQ4
ID   Q9HPQ4      PRELIMINARY;      PRT;   458 AA.
AC   Q9HPQ4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   VNG1524C.
GN   VNG1524C.
OS   Halobacterium sp. (strain NRC-1).
OC   Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
OC   Halobacterium.
OX   NCBI_TaxID=64091;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20504483; PubMed=11016950;
RA   Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA   Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA   Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA   Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA   Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA   Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA   Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA   Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT   "Genome sequence of Halobacterium species NRC-1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR   EMBL; AE005065; AAG19813.1; -.
DR   InterPro; IPR000954; Aminotran_3.
DR   InterPro; IPR001220; Lectin_legB.
DR   Pfam; PF00202; aminotran_3; 3.
DR   PROSITE; PS00600; AA_TRANSFER_CLASS_3; 1.
DR   PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
KW   Complete proteome.
SQ   SEQUENCE   458 AA;  49439 MW;  B1EA7132978E0F0E CRC64;
```

```
  Query Match             63.1%;  Score 41;  DB 17;  Length 458;
  Best Local Similarity   55.6%;  Pred. No. 16;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CDCRGDCFC 9
          | | |:| |
Db    198 CTCEGECSC 206
```

```
RESULT   8
Q96EB1
ID   Q96EB1      PRELIMINARY;      PRT;   535 AA.
AC   Q96EB1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SIMILAR TO HYPOTHETICAL PROTEIN.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS, AND EMBRYONAL CARCINOMA;
RA   Strausberg R.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC012514; AAH12514.1; -.
SQ   SEQUENCE   535 AA;  58713 MW;  86E6DD3B545E96D4 CRC64;
```

```
  Query Match             63.1%;  Score 41;  DB 4;  Length 535;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 CRGDCFC 9
          |:||| |
Db    522 CKGDCIC 528
```

```
RESULT   9
Q9SVX7
ID   Q9SVX7      PRELIMINARY;      PRT;   736 AA.
AC   Q9SVX7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 84.2 KDA PROTEIN.
GN   F15B8.180.
```

```
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Benes V., Rechmann S., Borkova D., Ansorge W., Mewes H.W.,
RA   Mayer K.F.X., Lemcke K., Schueller C., Quetier F., Salanoubat M.;
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL049660; CAB41192.1; -.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR002049; Laminin_EGF.
DR   PRINTS; PR00011; EGFLAMININ.
DR   SMART; SM00181; EGF; 2.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR   PROSITE; PS01186; EGF_2; 2.
KW   EGF-like domain; Glycoprotein; Hypothetical protein.
SQ   SEQUENCE   736 AA;  84202 MW;  349E0F1EE6A28C9A CRC64;


  Query Match             63.1%;  Score 41;  DB 10;  Length 736;
  Best Local Similarity   53.8%;  Pred. No. 24;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    4;  Gaps    1;

Qy       1 CDCRGDC----FC 9
           |||: ||    ||
Db     257 CDCKYDCLWGRFC 269


RESULT  10
P91904
ID   P91904      PRELIMINARY;      PRT;  3704 AA.
AC   P91904;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   LAMININ ALPHA (EPI-1 PROTEIN).
GN   EPI-1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Joh K., Hedgecock E.M.;
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Joh K., Zhu K., Hedgecock E.M., Inoue T., Hori K.;
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Berks M.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB001074; BAA19229.1; -.
DR   EMBL; AB016806; BAA32347.1; -.
DR   EMBL; Z70286; CAB61016.1; -.
DR   HSSP; P02468; 1KLO.
DR   InterPro; IPR001542; Arthro_defensin.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR000034; Laminin_B.
DR   InterPro; IPR002049; Laminin_EGF.
DR   InterPro; IPR001791; Laminin_G.
DR   InterPro; IPR001886; LamNT.
DR   InterPro; IPR001368; TNFR_c6.
DR   Pfam; PF00052; laminin_B; 1.
DR   Pfam; PF00053; laminin_EGF; 21.
DR   Pfam; PF00054; laminin_G; 5.
DR   Pfam; PF00055; laminin_Nterm; 1.
DR   PRINTS; PR00011; EGFLAMININ.
DR   ProDom; PD002082; LamNT; 1.
DR   ProDom; PD003031; Laminin_B; 1.
DR   SMART; SM00180; EGF_Lam; 21.
DR   SMART; SM00281; LamB; 1.
DR   SMART; SM00282; LamG; 5.
DR   SMART; SM00136; LamNT; 1.
DR   PROSITE; PS00425; ARTHROPOD_DEFENSINS; UNKNOWN_1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_19.
DR   PROSITE; PS01186; EGF_2; 4.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; 21.
DR   PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
KW   EGF-like domain; Glycoprotein; Laminin EGF-like domain; Repeat.
SQ   SEQUENCE   3704 AA;  407842 MW;  A2D5B66D7153919A CRC64;


  Query Match             63.1%;  Score 41;  DB 5;  Length 3704;
  Best Local Similarity   55.6%;  Pred. No. 99;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           ||  | |:|
Db     668 CDSNGQCYC 676


RESULT  11
Q962G0
ID   Q962G0      PRELIMINARY;      PRT;  100 AA.
AC   Q962G0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE METALLOTHIONEIN.
OS   Littorina littorea.
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC   Mesogastropoda; Littorinoidea; Littorinidae; Littorina.
OX   NCBI_TaxID=31216;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   English T.E., Storey K.B.;
RT   "Environmental stress-induced expression of a putative metallothionein
RT   gene in Littorina littorea.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY034179; AAK56498.1; -.
SQ   SEQUENCE   100 AA;  10039 MW;  085B08F60D91B040 CRC64;


  Query Match             61.5%;  Score 40;  DB 5;  Length 100;
  Best Local Similarity   55.6%;  Pred. No. 6.1;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           |:|: || |
Db      57 CNCKEDCRC 65


RESULT  12
Q9IBN0
ID   Q9IBN0      PRELIMINARY;      PRT;  116 AA.
AC   Q9IBN0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   POLYOMA VIRUS (NG-59) SMALL AND MIDDLE T ANTIGEN (FRAGMENT).
OS   Polyomavirus.
OC   Viruses; dsDNA viruses, no RNA stage; Polyomaviridae.
OX   NCBI_TaxID=10624;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=80072085; PubMed=6243123;
```

```
RA   Carmichael G.G., Benjamin T.L.;
RT   "Identification of DNA sequence changes leading to loss of
RT.  transforming ability in polyoma virus.";
RL   J. Biol. Chem. 255:230-235(1980).
DR   EMBL; K03531; AAA46896.1; -.
DR   InterPro; IPR003354; Papo_T_antigen.
DR   Pfam; PF02380; Papo_T_antigen; 1.
FT   NON_TER       1      1
FT   NON_TER     116    116
SQ   SEQUENCE   116 AA;  13658 MW;  10C8C62F899B51BF CRC64;


  Query Match             61.5%;  Score 40;  DB 12;  Length 116;
  Best Local Similarity   53.8%;  Pred. No. 7;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    4;  Gaps    1;

Qy      1 CDCR----GDCFC 9
          || |    |:|||
Db     62 CDARCLVLGECFC 74


RESULT  13
Q84251
ID   Q84251      PRELIMINARY;      PRT;   119 AA.
AC   Q84251;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   SMALL T ANTIGEN (FRAGMENT).
OS   Polyomavirus.
OC   Viruses; dsDNA viruses, no RNA stage; Polyomaviridae.
OX   NCBI_TaxID=10624;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=80072085; PubMed=6243123;
RA   Carmichael G.G., Benjamin T.L.;
RT   "Identification of DNA sequence changes leading to loss of
RT   transforming ability in polyoma virus.";
RL   J. Biol. Chem. 255:230-235(1980).
DR   EMBL; K03529; AAA46892.1; -.
DR   InterPro; IPR003354; Papo_T_antigen.
DR   Pfam; PF02380; Papo_T_antigen; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   119 AA;  14033 MW;  C8603B1391F3A134 CRC64;


  Query Match             61.5%;  Score 40;  DB 12;  Length 119;
  Best Local Similarity   53.8%;  Pred. No. 7.1;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    4;  Gaps    1;

Qy      1 CDCR----GDCFC 9
          || |    |:|||
Db     62 CDARCLVLGECFC 74


RESULT  14
Q84326
ID   Q84326      PRELIMINARY;      PRT;   167 AA.
AC   Q84326;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MIDDLE T ANTIGEN (FRAGMENT).
OS   Polyomavirus.
OC   Viruses; dsDNA viruses, no RNA stage; Polyomaviridae.
OX   NCBI_TaxID=10624;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=80072085; PubMed=6243123;
RA   Carmichael G.G., Benjamin T.L.;
RT   "Identification of DNA sequence changes leading to loss of
RT   transforming ability in polyoma virus.";
RL   J. Biol. Chem. 255:230-235(1980).
DR   EMBL; K03530; AAA46894.1; -.
DR   InterPro; IPR003354; Papo_T_antigen.
DR   Pfam; PF02380; Papo_T_antigen; 1.
FT   NON_TER       1      1
FT   NON_TER     167    167
SQ   SEQUENCE   167 AA;  19085 MW;  5B792F4D11E4C906 CRC64;


  Query Match             61.5%;  Score 40;  DB 12;  Length 167;
  Best Local Similarity   53.8%;  Pred. No. 9.6;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    4;  Gaps    1;

Qy      1 CDCR----GDCFC 9
          || |    |:|||
Db     15 CDARCLVLGECFC 27


RESULT  15
Q04190
ID   Q04190      PRELIMINARY;      PRT;   195 AA.
AC   Q04190;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SMALL T  ANTIGEN.
OS   Mus musculus (Mouse).
OG   Plasmid LFI.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91305109; PubMed=1649455;
RA   Yoshimura H., Ikeda Y., Yoshimoto M., Tamaki S., Hanada K., Kusano T.,
RA   Kohda T., Saito H., Oishi M.;
RT   "Structural and functional analysis of a polyoma-related mammalian
RT   plasmid (L factor). The enhancer activity and plasmid establishment.";
RL   Nucleic Acids Res. 19:3633-3639(1991).
DR   EMBL; X59849; CAA42512.1; -.
DR   InterPro; IPR001623; DnaJ_N.
DR   InterPro; IPR003354; Papo_T_antigen.
DR   Pfam; PF00226; DnaJ; 1.
DR   Pfam; PF02380; Papo_T_antigen; 1.
DR   SMART; SM00271; DnaJ; 1.
KW   Plasmid.
SQ   SEQUENCE   195 AA;  22783 MW;  6B3C29A11D29FDF3 CRC64;


  Query Match             61.5%;  Score 40;  DB 11;  Length 195;
  Best Local Similarity   53.8%;  Pred. No. 11;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    4;  Gaps    1;

Qy      1 CDCR----GDCFC 9
          || |    |:|||
Db    138 CDARCLVLGECFC 150


Search completed: May 29, 2002, 09:56:05
Job time: 239 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 29, 2002, 09:41:00 ; Search time 30.57 Seconds
(without alignments)
32.701 Million cell updates/sec

Title: US-09-734-628-1
Perfect score: 65
Sequence: 1 CDCRGDCFC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	65	100.0	9	16	AAR76200	Alphav/beta3 and a
2	65	100.0	9	19	AAW60289	Tumour homing pept
3	65	100.0	9	19	AAW56034	Chimeric adenoviru
4	65	100.0	9	20	AAY43233	RGD-containing pep
5	65	100.0	9	20	AAY48821	Membrane dipeptida
6	65	100.0	9	20	AAY42255	Synthetic RGD-4C p
7	65	100.0	9	20	AAW93626	NGR receptor bindi
8	65	100.0	9	21	AAB21701	Human breast tumou
9	65	100.0	9	21	AAB17346	Integrin-binding p
10	65	100.0	9	21	AAB17928	TPO-mimetic peptid
11	65	100.0	9	21	AAB17964	Integrin-binding p
12	65	100.0	9	21	AAY90211	Alphav integrin ta
13	65	100.0	9	21	AAY44970	RGD-4C targeting s
14	65	100.0	9	21	AAY54271	Alpha Vbeta-3 bind
15	65	100.0	9	22	AAE11044	RGD-containing pep
16	65	100.0	9	22	AAE06279	Tumour homing pept
17	65	100.0	9	22	AAB97086	Integrin-binding p
18	65	100.0	9	22	AAB20271	Peptide that speci
19	65	100.0	9	22	AAB50242	Enhanced infectivi
20	65	100.0	10	21	AAB21716	Human tumour-homin
21	65	100.0	10	22	AAE08561	RGD-4C peptide mot
22	65	100.0	11	16	AAR76194	Integrin binding p
23	65	100.0	11	18	AAW11184	Free peptide. Syn
24	65	100.0	11	19	AAW60299	Tumour homing pept
25	65	100.0	11	19	AAW57199	RGD-containing pep
26	65	100.0	11	21	AAY58860	Membrane binding e
27	65	100.0	11	21	AAY54273	Peptide inhibiting
28	65	100.0	11	22	AAE06294	Double cyclic homi
29	65	100.0	12	19	AAW56052	Chimeric adenoviru
30	65	100.0	12	20	AAW95410	Integrin-binding p
31	65	100.0	13	21	AAY90158	UPAR targeting seq
32	65	100.0	14	18	AAW19833	RGD peptide motif.
33	65	100.0	14	19	AAW56051	Chimeric adenoviru
34	65	100.0	15	19	AAW56040	Chimeric adenoviru
35	65	100.0	15	20	AAY43228	RGD-containing pep
36	65	100.0	15	21	AAY90167	UPAR targeting seq
37	65	100.0	15	21	AAY54272	Peptide inserted b
38	65	100.0	21	20	AAW96218	AlphavBeta3 integr
39	65	100.0	23	20	AAW96220	Modified Gene 10 3
40	65	100.0	24	19	AAW56044	Chimeric adenoviru
41	65	100.0	25	21	AAB21940	Homing antimicrobi
42	65	100.0	25	22	AAE06517	Homing pro-apoptot
43	65	100.0	26	21	AAB21937	Homing antimicrobi
44	65	100.0	26	22	AAE06516	Homing pro-apoptot
45	65	100.0	277	20	AAW82730	Adenovirus SCAR.RG

ALIGNMENTS

RESULT 1
AAR76200
ID AAR76200 standard; peptide; 9 AA.
XX
AC AAR76200;
XX
DT 24-JAN-1996 (first entry)
XX
DE Alphav/beta3 and alphav/beta5 integrin binding peptide #4.
XX
KW High affinity; integrin binding peptide; alpha5/beta1; alphav/beta5;
KW alphav/beta3; RGD; stable configuration; wound healing;
KW osteoclast attachemnt; bone; angiogenesis; metastasis; tumour;
KW smooth muscle cell migration.
XX
OS Synthetic.
XX
PN WO9514714-A1.
XX
PD 01-JUN-1995.
XX
PF 22-NOV-1994; 94WO-US13542.
XX
PR 04-AUG-1994; 94US-0286861.
PR 24-NOV-1993; 93US-0158001.
XX
PA (LJOL-) LA JOLLA CANCER RES FOUND.
XX
PI Koivunen E, Ruoslahti E;
XX
DR WPI; 1995-206899/27.
XX
PT High affinity integrin binding peptides - can be used to attach
PT cells to a substrate, inhibit the attachment of osteoclasts to bone,

PT promote wound healing, inhibit angiogenesis, metastasis of tumours
PT and migration of smooth muscle cells
XX
PS Claim 21; Page 62; 86pp; English.
XX
CC The sequences given in AAR76185-200 and AAR79073-94 are high affinity
CC integrin binding peptides which bind to various integrins. Peptides
CC which bind to alpha5/beta1 integrins contain the motifs given in
CC AAR76185-86 and peptides which bind to alphav/beta5 and alphav/beta3
CC integrins contain the motif given in AAR76187. Alphav/beta5 integrins
CC are also bound by RGD containing peptides. These peptides assume a
CC conformationally stabilised configuration which is due to the
CC formation of a disulphide bond, a peptide bond or a lactam bond.
CC These peptides may be used for isolating the complementary integrin
CC from a sample mixture by contacting them under ionic conditions to
CC allow binding of the integrin to the peptide and then separating the
CC integrin from the peptide. They can be used for attaching cells to
CC a substrate, by binding them to the substrate with the cell. The
CC peptides promote wound healing when applied locally and inhibit the
CC attachment of osteoclasts to bone. They inhibit angiogenesis,
CC metastasis of tumours and migration of smooth muscle cells.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 2
AAW60289
ID AAW60289 standard; peptide; 9 AA.
XX
AC AAW60289;
XX
DT 24-AUG-1998 (first entry)
XX
DE Tumour homing peptide of the invention.
XX
KW Tumour homing peptide; in vivo panning;
KW alpha-V-containing integrin binding motif; tumour.
XX
OS Unidentified.
XX
PN WO9810795-A2.
XX
PD 19-MAR-1998.
XX
PF 10-SEP-1997; 97WO-US16086.
XX
PR 10-SEP-1996; 96US-0710067.
XX
PA (BURN-) BURNHAM INST.
XX
PI Pasqualini R, Ruoslahti E;
XX
DR WPI; 1998-207151/18.
XX
PT Tumour homing molecules and their conjugates - useful for, e g.
PT directing linked moiety to tumour containing angiogenic vasculature
XX
PS Claim 6; Page 91; 105pp; English.
XX
CC The present peptide represents a tumour homing peptide, and is produced
CC by in vivo panning. The peptide has an alpha-V-containing integrin
CC binding motif, Arg-Gly-Asp (RGD). The in vivo panning comprises
CC administering a library of diverse peptides to a subject having a
CC tumour, collecting a sample of the tumour, identifying a peptide that
CC homes to the tumour, collecting a sample of normal tissue corresponding
CC to the tumour, and determining that the peptide that homes to the
CC tumour is not present in the normal tissue. The tumour homing peptide can
CC be linked to a moiety (e.g. doxorubicin), and used to direct the
CC moiety to a tumour.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 19; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 3
AAW56034
ID AAW56034 standard; peptide; 9 AA.
XX
AC AAW56034;
XX
DT 29-JUL-1998 (first entry)
XX
DE Chimeric adenovirus fiber protein non-native amino acid sequence 3.
XX
KW Chimeric; adenovirus; fiber protein; binding; targeting; coat protein;
KW constrained peptide motif; gene therapy; cancer; heart disease;
KW autoimmune disorder.
XX
OS Synthetic.
OS Mastadenovirus.
XX
PN WO9807865-A1.
XX
PD 26-FEB-1998.
XX
PF 21-AUG-1997; 97WO-US14719.
XX
PR 21-AUG-1996; 96US-0701124.
XX
PA (GENV-) GENVEC INC.
XX
PI Kovesdi I, Roelvink PW, Wickham TJ;
XX
DR WPI; 1998-169169/15.
XX
PT Chimeric adenovirus fibre proteins - containing non-native amino
PT acid sequence to provide for binding and entry into cells,
PT especially for gene therapy
XX
PS Claim 7; Page 68; 124pp; English.
XX
CC The present sequence represents a specifically claimed non-native amino
CC acid sequence from a chimeric adenovirus fibre protein (AFP) of the
CC present invention. The non-native amino acid sequence allows the
CC chimeric fibre (or a vector comprising the chimeric fibre) to more
CC efficiently bind to and enter cells. The products can be used for gene
CC therapy, for treating cancer, e.g. melanoma, glioma and lung cancers as
CC well as genetic disorders, e.g. cystic fibrosis, haemophilia and
CC muscular dystrophy as well as pathogenic infections, e.g. HIV,
CC tuberculosis and hepatitis and also for heart disease, to e.g. prevent
CC restenosis following angioplasty or to promote angiogenesis to reperfuse
CC necrotic tissue, and in autoimmune disorders, e.g. Crohn's disease,
CC colitis, rheumatoid arthritis, and Alzheimer's disease.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 19; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;

```
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CDCRGDCFC 9
          |||||||||
Db      1 cdcrgdcfc 9


RESULT   4
AAY43233
ID   AAY43233 standard; peptide; 9 AA.
XX
AC   AAY43233;
XX
DT   13-JAN-2000  (first entry)
XX
DE   RGD-containing peptide #12.
XX
KW   Nucleic acid delivery vehicle; bifunctional complex; transgene; CFTR;
KW   cell surface targeting; cell surface molecule binding region; integrin;
KW   cystic fibrosis transmembrane regulator; alpha1-antitrypsin;
KW   suicide gene; beta-glucocerebrosidase; cell transfection; cell infection;
KW   RGD peptide.
XX
OS   Synthetic.
XX
PN   WO9940214-A2.
XX
PD   12-AUG-1999.
XX
PF   08-FEB-1999;   99WO-US02680.
XX
PR   09-FEB-1998;   98US-0020483.
PR   06-NOV-1998;   98US-0107471.
XX
PA   (GENZ ) GENZYME CORP.
XX
PI   O'riordan C,  Romanczuk H,  Wadsworth SC;
XX
DR   WPI; 1999-610583/52.
XX
PT   Nucleic acid delivery vehicles useful for transfecting and infecting a
PT   target cell  -
XX
PS   Claim 22; Page 39; 118pp; English.
XX
CC   This sequence represents a RGD-containing peptide that can be used in a
CC   bifunctional complex used in the nucleic acid delivery vehicle (I) of the
CC   invention. (I) is for transfecting and/or infecting a target cell, and
CC   comprises a transgene and a bifunctional complex (B) that targets the
CC   nucleic acid delivery vehicle to the cell surface. (B) comprises a
CC   delivery vehicle binding portion, a cell surface molecule binding portion
CC   (such as this sequence) and a linker connecting them. The delivery
CC   vehicle can be specifically targeted to the cell via the binding to cell
CC   surface molecules. (I) can be used to target cells, which express
CC   integrins such as, HT-29 colon carcinoma cells, lymphocytes and
CC   monocytes, blood platelets, SMC-90 human lung fibroblast, MG(63)
CC   osteosarcoma cell line, vascular endothelial cells and melanoma cells.
CC   (I) is useful for delivery of nucleic acids encoding CFTR (cystic
CC   fibrosis transmembrane regulator), alpha1-antitrypsin,
CC   beta-glucocerebrosidase and suicide genes. The construct increases the
CC   efficiency of cellular uptake of (I). The constructs also enable the
CC   transfection/infection of cells that are normally refractory to
CC   transfection/infection by targeting cell receptors that are present on
CC   such cells.
XX
SQ   Sequence   9 AA;


  Query Match           100.0%;  Score 65;  DB 20;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 6.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CDCRGDCFC 9
          |||||||||
Db      1 cdcrgdcfc 9


RESULT   5
AAY48821
ID   AAY48821 standard; Peptide; 9 AA.
XX
AC   AAY48821;
XX
DT   10-DEC-1999  (first entry)
XX
DE   Membrane dipeptidase-binding retina homing peptide #7.
XX
KW   Homing peptide; organ; tissue; lung; pancreas; skin; retina; MDP;
KW   prostate; ovary; lymph node; adrenal gland; liver; gut; tumour;
KW   membrane dipeptidase.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9946284-A2.
XX
PD   16-SEP-1999.
XX
PF   10-MAR-1999;   99WO-US05284.
XX
PR   13-MAR-1998;   98US-0042107.
PR   26-FEB-1999;   99US-0042107.
XX
PA   (BURN-) BURNHAM INST.
XX
PI   Rajotte D,  Pasqualini R,  Ruoslahti EI;
XX
DR   WPI; 1999-571717/48.
XX
PT   New peptides which selectively home to organs or tissues, used for,
PT   e.g. identifying target ligands and for therapy of pathological
PT   conditions  -
XX
PS   Example 6; Page 149; 193pp; English.
XX
CC   The present invention describes peptides that selectively home to a
CC   tissue or organ. The peptides can be used for identifying an organ
CC   or tissue, for identifying a target molecule expressed by an organ or
CC   tissue or for treating an organ or tissue pathology, where the organ or
CC   tissue is selected from prostate, lung, skin, retina, pancreas, gut,
CC   ovary, adrenal gland, liver, and lymph node. The peptide bind to the
CC   membrane dipeptidase (MDP). AAY48618 to AAY49066 represent sequences
CC   which are used in the exemplification of the present invention.
XX
SQ   Sequence   9 AA;


  Query Match           100.0%;  Score 65;  DB 20;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 6.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CDCRGDCFC 9
          |||||||||
Db      1 cdcrgdcfc 9


RESULT   6
AAY42255
ID   AAY42255 standard; peptide; 9 AA.
XX
AC   AAY42255;
XX
DT   01-DEC-1999  (first entry)
XX
DE   Synthetic RGD-4C peptide.
XX
```

KW Adenovirus; gene therapy; coxsackievirus adenovirus receptor;
KW CAR; cancer; cystic fibrosis; muscular dystrophy.
XX
OS Synthetic.
XX
PN WO9939734-A1.
XX
PD 12-AUG-1999.
XX
PF 05-FEB-1999; 99WO-US02549.
XX
PR 06-FEB-1998; 98US-0073947.
PR 10-SEP-1998; 98US-0099801.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Curiel DT, Krasnykh VN, Dmitriev I;
XX
DR WPI; 1999-539951/45.
XX
Recombinant adenovirus vectors with modified fiber knob loops, useful in gene therapy -
XX
PS Example 21; Page 49; 126pp; English.
XX
CC This sequence represents a synthetic RGD-4C peptide. DNA encoding
CC this sequence was cloned into the sequence encoding the HI loop of the
CC adenovirus fibre protein knob domain. This was then used in the
CC construction of plasmids encoding a modified fibre protein. Recombinant
CC adenovirus genomes were generated by homologous DNA recombination in E.
CC coli, before excision of the newly generated genome for virus rescue.
CC The knob domain of the adenovirus fibre protein mediates the initial
CC binding and recognition of the coxsackievirus and adenovirus receptor
CC (CAR) on the cell surface. The HI loop protrudes from the knob domain
CC and connects beta-strands involved in the formation of the cell binding
CC site. Recombinant adenovirus vectors are used in a number of gene
CC therapy applications; however, the reliance on the CAR means that
CC in certain situations, recombinant virions are sequestered by high
CC CAR-expressing non-target cells while the true target cells, if low
CC in CAR, receive little of the therapeutic gene. Modification of the HI
CC loop by replacement of the hypervariable region of the loop with a
CC peptide such as the RGD peptide results in the
CC ability of the virus to utilise an alternative receptor during the cell
CC entry process. Modifying the adenovirus fibre knob protein in this way
CC increases the ability of an adenovirus to transduce a tumour cell in
CC vitro, in vivo and ex vivo. The vector Ad5FHIFLAG incorporating an RGD
CC peptide demonstrated two to three orders of magnitude
CC of increased gene transfer to ovarian cancer cells. The modified
CC adenovirus has an altered tropism, which allows the adenovirus to be
CC targeted to selected cell types. The recombinant adenovirus can be used
CC to provide gene therapy for individuals suffering from cancer, cystic
CC fibrosis and Duchenne's muscular dystrophy.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 7
AAW93626
ID AAW93626 standard; Protein; 9 AA.
XX
AC AAW93626;
XX
DT 28-JUN-1999 (first entry)
XX
DE NGR receptor binding tumour homing peptide 5.
XX
KW Tumour homing peptide; tumour; diagnosis; endothelial cell;
KW angiogenic vasculature; anti-tumour; anti-inflammatory; anti-angiogenic;
KW anti-arthritic; NGR receptor; inhibitor; angiogenesis; anticancer drug;
KW prognosis; inflammation; regeneration; wounded tissue; targetting;
KW macular degeneration; diabetic retinopathy; rheumatoid arthritis;
KW occlusive thrombus.
XX
OS Synthetic.
XX
PN WO9913329-A1.
XX
PD 18-MAR-1999.
XX
PF 08-SEP-1998; 98WO-US18895.
XX
PR 25-AUG-1998; 98US-0139802.
PR 10-SEP-1997; 97US-0926914.
XX
PA (BURN-) BURNHAM INST.
XX
PI Pasqualini R, Ruoslahti E;
XX
DR WPI; 1999-215158/18.
XX
PT Identifying molecules that home to angiogenic vasculature used as
PT targets for anticancer agents
XX
PS Claim 15; Page 7; 180pp; English.
XX
CC This invention describes novel peptides which home to angiogenic
CC vasculature, specifically of a tumour and which have anti-tumour,
CC anti-inflammatory, anti-angiogenic and anti-arthritic activity. Such
CC molecules are identified by treating a purified NGR receptor with a test
CC compound and identifying compounds that bind specifically to the NGR
CC receptor. The peptides of the invention are inhibitors of angiogenesis
CC and can be used to produce conjugates for delivering agents to
CC angiogenic vasculature, particularly anticancer drugs or an imaging
CC agent, for diagnosis or prognosis. These conjugates may be directed to
CC non-tumour angiogenic vasculature, e.g. that present in inflammatory,
CC regenerating or wounded tissue, e.g. for treatment of macular
CC degeneration, diabetic retinopathy or rheumatoid arthritis. The peptides
CC provide specific targeting to tumours, especially their supporting
CC vasculature, since the NGR receptor is exposed to the circulation only in
CC angiogenic vasculature. Precise targeting should reduce the systemic
CC toxicity of anticancer drugs in the conjugates. Complete killing of all
CC target cells may not be essential since partial denudation of endothelium
CC may result in an occlusive thrombus, and endothelial cells are unlikely
CC to become resistant to anticancer agents nor to lose the targeting
CC receptor. AAW93622-W93809 and AAW93843-44 are examples of tumour homing
CC peptides used in the invention.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 8
AAB21701
ID AAB21701 standard; Peptide; 9 AA.
XX
AC AAB21701;
XX
DT 22-MAR-2001 (first entry)
XX

DE Human breast tumour homing peptide #1.
XX
KW Cytostatic; homing pro-apoptotic conjugate; tumour; antimicrobial;
KW breast; prostate; melanoma; cancer; Kaposi's sarcoma; human.
XX
OS Homo sapiens.
XX
PN WO200042973-A2.
XX
PD 27-JUL-2000.
XX
PF 21-JAN-2000; 2000WO-US01602.
XX
PR 22-JAN-1999; 99US-0235902.
XX
PA (BURN-) BURNHAM INST.
XX
PI Ellerby HM, Bredesen DE, Pasqualini R, Ruoslahti EI;
XX
DR WPI; 2000-499174/44.
XX
PT Homing pro-apoptotic conjugate comprising a tumor homing molecule that
PT selectively homes to a mammalian cell type or tissue linked to an
PT antimicrobial peptide, useful for the treatment of prostate cancer -
XX
PS Claim 12; Page 105; 118pp; English.
XX
CC The present invention relates to homing pro-apoptotic conjugates,
CC comprising of a tumour homing molecule that selectively homes to a
CC mammalian cell type or tissue, linked to an antimicrobial peptide. The
CC homing pro-apoptotic conjugates are selectively internalised by the
CC mammalian cell type or tissue and exhibits high toxicity, especially to
CC angiogenic vasculature. The antimicrobial peptide has low mammalian cell
CC toxicity when not linked to the tumor homing molecule. The conjugates are
CC useful for the treatment of cancer e.g. Kaposi's sarcoma, breast and
CC prostate cancer or melanoma. The present sequence is a homing peptide
CC isolated in the present invention, which can be conjugated to an
CC antimicrobial peptide to make the homing pro-apoptotic conjugates of the
CC present invention.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 9
AAB17346
ID AAB17346 standard; Peptide; 9 AA.
XX
AC AAB17346;
XX
DT 31-OCT-2000 (first entry)
XX
DE Integrin-binding peptide sequence SEQ ID NO:450.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Claim 39; Page 354; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 10
AAB17928
ID AAB17928 standard; Peptide; 9 AA.
XX
AC AAB17928;
XX
DT 31-OCT-2000 (first entry)
XX
DE TPO-mimetic peptide sequence SEQ ID NO:1032.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX

PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Disclosure; Page 559; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 11
AAB17964
ID AAB17964 standard; Peptide; 9 AA.
XX
AC AAB17964;
XX
DT 31-OCT-2000 (first entry)
XX
DE Integrin-binding peptide sequence SEQ ID NO:1076.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Claim 39; Page 591; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 12
AAY90211
ID AAY90211 standard; peptide; 9 AA.
XX
AC AAY90211;
XX
DT 21-SEP-2000 (first entry)
XX
DE Alphav integrin targeting peptide #1.
XX
KW Ligand epitope; UPAR; urokinase-type plasminogen activator receptor;
KW adenovirus; hexon HVR5 loop; hexon HI loop; peripheral artery disease;
KW recombinant adenovirus vector; tumour; restenosis; gene therapy; asthma;
KW smooth muscle cell proliferation inhibitor; coronary artery disease;
KW obesity; neurodegenerative disease; infection; autoimmune disease; HIV;
KW thrombosis; diabetes; tropism-modified virus.
XX
OS Adenovirus sp.
XX
PN WO200012738-A1.
XX
PD 09-MAR-2000.
XX
PF 27-AUG-1999; 99WO-IB01524.

XX
PR 27-AUG-1998; 98US-0098028.
XX
PA (AVET) AVENTIS PHARMA SA.
XX
PI Vigne E, Dedieu J, Latta M, Yeh P, Perricaudet M;
XX
DR WPI; 2000-256653/22.
XX
PT Urokinase-type plasminogen activator receptor (UPAR)-targeted
PT adenovirus vectors having modified hexon HVR5 and HI loops and modified
PT fiber proteins useful for targeted gene therapy to treat cancer or
PT restenosis -
XX
PS Example 5; Page 53; 128pp; English.
XX
CC This sequence represents a alphav integrin targeting peptide.
CC The invention relates to an adenovirus from which at
CC least a part of the hexon HVR5 or HI loop is replaced with a binding
CC peptide, or targeting sequence, flanked by connecting amino acid spacers,
CC to functionally display its binding specificity at the capsid surface.
CC The invention also relates to a recombinant adenovirus vector where a
CC binding peptide, or targeting sequence, is connected to the C-terminus of
CC the fiber by a connecting spacer, or linker, so as to functionally
CC display its binding specificity at the capsid surface. The adenovirus or
CC recombinant adenovirus vector can be used to preferentially express a
CC gene in a target cell, especially a cell that expresses a UPAR. The
CC targeted adenovirus vector preferably comprises a heterologous gene
CC encoding a gene for treatment of a tumour or restenosis. The targeted
CC adenovirus vector is useful for gene therapy treatment of a disease, and
CC for manufacturing a medicine used in gene therapy treatment of a disease.
CC The viruses can also be used to inhibit smooth muscle cell proliferation,
CC to treat peripheral artery diseases, coronary artery diseases, obesity,
CC neurodegenerative diseases, infections, autoimmune diseases, asthma, HIV,
CC thrombosis, and diabetes. The viruses are particularly targeted against a
CC urokinase-type plasminogen activator receptor (UPAR). The adenoviruses
CC are tropism-modified without adversely impacting productivity of the
CC vectors.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 13
AAY44970
ID AAY44970 standard; Protein; 9 AA.
XX
AC AAY44970;
XX
DT 23-MAY-2000 (first entry)
XX
DE RGD-4C targeting sequence for KDEL receptor inhibitor protein.
XX
KW KDEL receptor inhibitor; heat shock protein; immune response;
KW oligomerisation domain; neoplasia; sarcoma; lymphoma; leukaemia;
KW melanoma; carcinoma; glioblastoma; astromytoma; oncogene;
KW infectious disease; allergy; autoimmune disease.
XX
OS Unidentified.
XX
PN WO200006729-A1.
XX
PD 10-FEB-2000.
XX
PF 28-JUL-1999; 99WO-US17147.
XX
PR 29-JUL-1998; 98US-0124671.
XX
PA (SLOK) SLOAN KETTERING INST CANCER RES.
XX
PI Rothman JE, Mayhew M, Hoe MH;
XX
DR WPI; 2000-195296/17.
XX
PT Inhibitors of the KDEL receptor which comprises an oligomerization
PT domain useful for promoting secretion of proteins which are normally
PT retained within the cell -
XX
PS Disclosure; Page 17; 87pp; English.
XX
CC The patent discloses the use of KDEL receptor inhibitor to promote
CC secretion of proteins that are normally retained within the cell such as
CC heat shock proteins by inhibiting KDEL receptor-mediated return of
CC protein complexes to endoplasmic reticulum. This makes the secreted heat
CC shock proteins more accessible to the immune system and improves immune
CC response to a target antigen. The inhibitor protein comprises several
CC subunits where each subunit comprises an oligomerisation domain and has
CC at its carboxy terminus a region which binds to a KDEL receptor. The
CC target antigen may be associated with diseases including neoplasia such
CC as sarcoma, lymphoma, leukemia, melanoma, carcinoma, glioblastoma and
CC astromytoma, with defective tumour suppressor genes, oncogenes,
CC infectious diseases, allergy or autoimmune diseases. The present
CC sequence is a targeting peptide termed RGD-4C. This may be incorporated
CC into the amino terminal region of a KDEL receptor inhibitor protein
CC downstream from a cleavably removed sequence to improve its activity or
CC alter its immunogenicity.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 14
AAY54271
ID AAY54271 standard; Peptide; 9 AA.
XX
AC AAY54271;
XX
DT 06-APR-2000 (first entry)
XX
DE Alpha Vbeta-3 binding peptide sequence.
XX
KW Envelope protein; mutant; retrovirus; surface protein shedding;
KW envelope protein stability; gene therapy; drug therapy; cancer;
KW adenosine deaminase deficiency; thalassemia; hemophilia; diabetes;
KW alpha-anti trypsin deficiency; brain disorder; neural disorder;
KW phenylketonuria; growth disorder; heart disease; immune disease.
XX
OS Unidentified.
XX
PN WO9960110-A2.
XX
PD 25-NOV-1999.
XX
PF 20-MAY-1999; 99WO-US11155.
XX
PR 20-MAY-1998; 98US-0086149.
XX
PA (UYTE-) UNIV TENNESSEE RES CORP.
XX
PI Albritton LM, Zavorotinskaya T;

XX
DR WPI; 2000-116313/10.
XX
PT Novel isolated nucleic acid, useful for gene therapy -
XX
PS Example 10; Page 84; 190pp; English.
XX
CC The specification describes mutant retrovirus envelope proteins. The
CC envelope protein coding sequence can be mutated to encode a mutant
CC envelope protein with a substitution of one or more amino acids in at
CC least one motif of the retrovirus protein. The mutant protein fragment
CC allows for decreased shedding of the surface protein by suppressing
CC precursor cleavage and increase envelope stability and fusion of
CC retroviruses with cell membranes, while maintaining mutant envelope
CC protein incorporation into a virion, and viral titers of about two orders
CC of magnitude within that observed for wild-type retrovirus when the
CC protein or fragment is expressed on the surface of a retroviral particle.
CC The proteins have an increased ability to penetrate targets, typically
CC cells and a correspondingly increased ability to deliver nucleic acids or
CC drugs. The mutated nucleic acid is useful for gene and drug therapy,
CC especially as drug delivery vehicles. The retrovirus particles can be
CC utilized to transduce eukaryotic cells. The transduced cells are useful
CC in the treatment of cancer in a human. Other diseases contemplated for
CC treatment include adenosine deaminase deficiency (ADA), thalassemia,
CC hemophilia, diabetes, alpha-anti trypsin deficiency, brain and neural
CC disorders, phenylketonuria, growth disorders, heart diseases and immune
CC diseases. The present sequence was used in the course of the invention,
CC to quantitate targeted retroviral vector gene delivery in vivo.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 15
AAE11044
ID AAE11044 standard; peptide; 9 AA.
XX
AC AAE11044;
XX
DT 18-DEC-2001 (first entry)
XX
RGD-containing peptide.
XX
KW Tumour necrosis factor; TNF; cytokine; cytostatic; virucide;
KW TNF related apoptosis inducing ligand; TRAIL; cancer; viral infection;
KW human immunodeficiency virus; HIV; leukaemia; gene therapy; lymphoma;
KW melanoma.
XX
OS Unidentified.
XX
PN US6284236-B1.
XX
PD 04-SEP-2001.
XX
PF 26-MAY-1999; 99US-0320424.
XX
PR 29-JUN-1995; 95US-0496632.
PR 01-NOV-1995; 95US-0548368.
PR 25-JUN-1996; 96US-0670354.
PR 26-MAR-1998; 98US-0048641.
PR 10-NOV-1998; 98US-0190046.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Wiley SR, Goodwin RG;

XX
DR WPI; 2001-595463/67.
XX
PT New tumor necrosis factor related apoptosis inducing ligand
PT polypeptides for treating viral infections (e.g. bovine viral diarrhea
PT or human immunodeficiency virus), or cancers (e.g. leukemia or
PT lymphoma) -
XX
PS Disclosure; Column 11; 41pp; English.
XX
CC The invention relates to a cytokine designated as tumour necrosis
CC factor (TNF) related apoptosis inducing ligand (TRAIL), which induces
CC apoptosis of certain target cells, including cancer cells and virally
CC infected cells. The TRAIL polypeptides are useful in killing cancer
CC cells, in treating viral infections (e.g. bovine viral diarrhoea or
CC human immunodeficiency virus (HIV)) and cancers (e.g. leukaemia,
CC lymphoma and melanoma), as a research reagent useful in studying
CC apoptosis including the regulation of programmed cell death. TRAIL
CC DNA sequences may be employed in developing a gene therapy approach
CC to treating disorders mediated by defective or insufficient amounts
CC of TRAIL, in the production of TRAIL polypeptides and as probes or
CC primers in polymerase chain reactions (PCR). The present sequence is
CC a RGD-containing peptide that binds an integrin associated with
CC tumour. This sequence is used to construct a fusion protein
CC comprising TRAIL protein.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

Search completed: May 29, 2002, 09:52:03
Job time: 663 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 29, 2002, 10:42:42 ; Search time 12.96 Seconds
(without alignments)
16.962 Million cell updates/sec

Title: US-09-734-628-1
Perfect score: 65
Sequence: 1 CDCRGDCFC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	65	100.0	9	2	US-08-701-124-3	Sequence 3, Appli
2	65	100.0	9	2	US-08-286-861-16	Sequence 16, Appl
3	65	100.0	9	3	US-09-026-633-1	Sequence 1, Appli
4	65	100.0	9	3	US-09-130-225-3	Sequence 3, Appli
5	65	100.0	9	4	US-09-124-671-33	Sequence 33, Appl
6	65	100.0	9	4	US-09-258-754-211	Sequence 211, App
7	65	100.0	9	4	US-09-139-802-1	Sequence 1, Appli
8	65	100.0	9	4	US-09-042-107-211	Sequence 211, App
9	65	100.0	9	4	US-09-320-424-20	Sequence 20, Appl
10	65	100.0	9	4	US-09-426-680-12	Sequence 12, Appl
11	65	100.0	9	4	US-09-455-061-3	Sequence 3, Appli
12	65	100.0	11	2	US-08-717-169-17	Sequence 17, Appl
13	65	100.0	11	2	US-08-286-861-10	Sequence 10, Appl
14	65	100.0	11	4	US-09-139-802-16	Sequence 16, Appl
15	65	100.0	12	2	US-08-701-124-79	Sequence 79, Appl
16	65	100.0	12	3	US-09-130-225-79	Sequence 79, Appl
17	65	100.0	12	4	US-09-455-061-79	Sequence 79, Appl
18	65	100.0	14	2	US-08-701-124-68	Sequence 68, Appl
19	65	100.0	14	3	US-09-130-225-68	Sequence 68, Appl
20	65	100.0	14	4	US-09-455-061-68	Sequence 68, Appl
21	65	100.0	15	2	US-08-701-124-31	Sequence 31, Appl
22	65	100.0	15	3	US-09-130-225-31	Sequence 31, Appl
23	65	100.0	15	4	US-09-426-680-7	Sequence 7, Appli
24	65	100.0	15	4	US-09-455-061-31	Sequence 31, Appl
25	65	100.0	24	2	US-08-701-124-49	Sequence 49, Appl
26	65	100.0	24	3	US-09-130-225-49	Sequence 49, Appl
27	65	100.0	24	4	US-09-455-061-49	Sequence 49, Appl
28	59	90.8	9	2	US-08-286-861-17	Sequence 17, Appl
29	56	86.2	8	3	US-09-026-633-4	Sequence 4, Appli
30	56	86.2	13	4	US-09-426-680-8	Sequence 8, Appli
31	51	78.5	9	2	US-08-701-124-4	Sequence 4, Appli
32	51	78.5	9	2	US-08-286-861-15	Sequence 15, Appl
33	51	78.5	9	3	US-09-130-225-4	Sequence 4, Appli
34	51	78.5	9	4	US-09-455-061-4	Sequence 4, Appli
35	49	75.4	9	2	US-08-286-861-18	Sequence 18, Appl
36	44	67.7	7	4	US-09-426-680-11	Sequence 11, Appl
37	44	67.7	577	2	US-07-728-215-29	Sequence 29, Appl
38	44	67.7	577	4	US-08-938-085A-29	Sequence 29, Appl
39	44	67.7	788	2	US-07-728-215-27	Sequence 27, Appl
40	44	67.7	788	4	US-08-938-085A-27	Sequence 27, Appl
41	40	61.5	8	1	US-08-421-702A-22	Sequence 22, Appl
42	40	61.5	8	1	US-08-303-052A-22	Sequence 22, Appl
43	40	61.5	8	1	US-08-421-696A-22	Sequence 22, Appl
44	40	61.5	8	1	US-08-421-697A-22	Sequence 22, Appl
45	40	61.5	8	1	US-08-421-698A-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-08-701-124-3
; Sequence 3, Application US/08701124
; Patent No. 5846782
; GENERAL INFORMATION:
; APPLICANT: Wickham, Thomas J.
; APPLICANT: Roelvink, Petrus W.
; APPLICANT: Kovesdi, Imre
; TITLE OF INVENTION: TARGETING ADENOVIRUS WITH USE OF
; TITLE OF INVENTION: CONSTRAINED PEPTIDE MOTIFS
; NUMBER OF SEQUENCES: 80
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
; STREET: Two Prudential Plaza - 49th Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/701,124
; FILING DATE: 21-AUG-1996
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-701-124-3

Query Match 100.0%; Score 65; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 CDCRGDCFC 9

RESULT 2
US-08-286-861-16
; Sequence 16, Application US/08286861
; Patent No. 5981478
; GENERAL INFORMATION:

```
;    APPLICANT:  Ruoslahti, Erkki
;    APPLICANT:  Koivunen, Erkki
;    TITLE OF INVENTION:  No. 5981478el Integrin-Binding Peptides
;    NUMBER OF SEQUENCES:  46
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/286,861
;      FILING DATE:  04-AUG-1994
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/158,001
;      FILING DATE:  24-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-LA 9992
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  circular
US-08-286-861-16


  Query Match           100.0%;  Score 65;  DB 2;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CDCRGDCFC 9
          |||||||||
Db      1 CDCRGDCFC 9


RESULT   3
US-09-026-633-1
; Sequence 1, Application US/09026633
; Patent No. 6025328
; GENERAL INFORMATION:
;  APPLICANT: McMorris, Trevor C.
;  APPLICANT: Kelner, Michael J.
;  TITLE OF INVENTION: Antitumor agents
;  FILE REFERENCE: 103.008US1
;  CURRENT APPLICATION NUMBER: US/09/026,633
;  CURRENT FILING DATE: 1998-02-20
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Amino acid sequence
US-09-026-633-1


  Query Match           100.0%;  Score 65;  DB 3;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CDCRGDCFC 9
          |||||||||
Db      1 CDCRGDCFC 9


RESULT   4
US-09-130-225-3
; Sequence 3, Application US/09130225
; Patent No. 6057155
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Roelvink, Petrus W.
;    APPLICANT:  Kovesdi, Imre
;    TITLE OF INVENTION:  TARGETING ADENOVIRUS WITH USE OF
;    TITLE OF INVENTION:  CONSTRAINED PEPTIDE MOTIFS
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza - 49th Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/130,225
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 8-701124
;      FILING DATE:  21-AUG-1996
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-130-225-3


  Query Match           100.0%;  Score 65;  DB 3;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 1.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CDCRGDCFC 9
          |||||||||
Db      1 CDCRGDCFC 9


RESULT   5
US-09-124-671-33
; Sequence 33, Application US/09124671A
; Patent No. 6160088
; GENERAL INFORMATION:
;  APPLICANT: Rothman, James
;  APPLICANT: Mayhew, Mark
;  APPLICANT: Hoe, Mee
;  TITLE OF INVENTION: KDEL RECEPTOR INHIBITORS
;  FILE REFERENCE: 31488
;  CURRENT APPLICATION NUMBER: US/09/124,671A
;  CURRENT FILING DATE: 1998-07-29
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 33
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: alpha-five integrin binding motif
```

US-09-124-671-33

Query Match 100.0%; Score 65; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 CDCRGDCFC 9

RESULT 6
US-09-258-754-211
; Sequence 211, Application US/09258754
; Patent No. 6174687
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; APPLICANT: Rajotte, Daniel
; TITLE OF INVENTION: Methods of Identifying Lung Homing Molecules Using
; TITLE OF INVENTION: Membrane Dipeptidase
; FILE REFERENCE: P-LJ 3443
; CURRENT APPLICATION NUMBER: US/09/258,754
; CURRENT FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 09/042,107
; EARLIER FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 452
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 211
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-258-754-211

Query Match 100.0%; Score 65; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 CDCRGDCFC 9

RESULT 7
US-09-139-802-1
; Sequence 1, Application US/09139802
; Patent No. 6180084
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: NGR Receptor and Methods of Identifying Tumor Homing
; TITLE OF INVENTION: Molecules That Home to Angiogenic Vasculature Using
; TITLE OF INVENTION: Same
; FILE REFERENCE: P-LJ 3203
; CURRENT APPLICATION NUMBER: US/09/139,802
; CURRENT FILING DATE: 1998-08-25
; EARLIER APPLICATION NUMBER: 08/926,914
; EARLIER FILING DATE: 1997-09-10
; EARLIER APPLICATION NUMBER: 08/710,067
; EARLIER FILING DATE: 1996-09-10
; NUMBER OF SEQ ID NOS: 226
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide

US-09-139-802-1

Query Match 100.0%; Score 65; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 CDCRGDCFC 9

RESULT 8
US-09-042-107-211
; Sequence 211, Application US/09042107
; Patent No. 6232287
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Molecules that Home to Various Selected Organs or
; TITLE OF INVENTION: Tissues
; FILE REFERENCE: P-LJ 2892
; CURRENT APPLICATION NUMBER: US/09/042,107
; CURRENT FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 436
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 211
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-042-107-211

Query Match 100.0%; Score 65; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 CDCRGDCFC 9

RESULT 9
US-09-320-424-20
; Sequence 20, Application US/09320424
; Patent No. 6284236
; GENERAL INFORMATION:
; APPLICANT: Wiley, Steven R.
; APPLICANT: Goodwin, Raymond G.
; TITLE OF INVENTION: Cytokine that Induces Apoptosis
; FILE REFERENCE: 2835-E
; CURRENT APPLICATION NUMBER: US/09/320,424
; CURRENT FILING DATE: 1999-05-26
; EARLIER APPLICATION NUMBER: 09/190,046
; EARLIER FILING DATE: 1998-11-10
; EARLIER APPLICATION NUMBER: 09/048,641
; EARLIER FILING DATE: 1998-03-26
; EARLIER APPLICATION NUMBER: 08/670,354
; EARLIER FILING DATE: 1996-06-25
; EARLIER APPLICATION NUMBER: 08/548,368
; EARLIER FILING DATE: 1995-11-01
; EARLIER APPLICATION NUMBER: 08/496,632
; EARLIER FILING DATE: 1995-06-29
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: artificial

```
;   OTHER INFORMATION: peptide
US-09-320-424-20


  Query Match             100.0%;  Score 65;  DB 4;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           |||||||||
Db       1 CDCRGDCFC 9


RESULT  10
US-09-426-680-12
; Sequence 12, Application US/09426680
; Patent No. 6287857
; GENERAL INFORMATION:
;  APPLICANT: Catherine R. O'Riordan
;  APPLICANT: Samuel C. Wadsworth
;  TITLE OF INVENTION: Nucleic Acid Delivery Vehicles
;  FILE REFERENCE: GA0103USB2
;  CURRENT APPLICATION NUMBER: US/09/426,680
;  CURRENT FILING DATE: 1999-10-25
;  EARLIER APPLICATION NUMBER: PCT/US99/02680
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 12
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: human
;   FEATURE:
;   NAME/KEY: PEPTIDE
;   LOCATION: (0)...(0)
US-09-426-680-12


  Query Match             100.0%;  Score 65;  DB 4;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           |||||||||
Db       1 CDCRGDCFC 9


RESULT  11
US-09-455-061-3
; Sequence 3, Application US/09455061
; Patent No. 6329190
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Roelvink, Petrus W.
;    APPLICANT:  Kovesdi, Imre
;    TITLE OF INVENTION:  TARGETING ADENOVIRUS WITH USE OF
;    TITLE OF INVENTION:  CONSTRAINED PEPTIDE MOTIFS
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza - 49th Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/455,061
;      FILING DATE:  06-DEC-1999
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 9-130225
;      FILING DATE:  06-AUG-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 8-701124
;      FILING DATE:  21-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hefner, M. Daniel
;      REGISTRATION NUMBER:  41,826
;      REFERENCE/DOCKET NUMBER:  203128
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-455-061-3


  Query Match             100.0%;  Score 65;  DB 4;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           |||||||||
Db       1 CDCRGDCFC 9


RESULT  12
US-08-717-169-17
; Sequence 17, Application US/08717169
; Patent No. 5922676
;  GENERAL INFORMATION:
;    APPLICANT:  Pasqualini, Renata
;    APPLICANT:  Ruoslahti, Erkki
;    TITLE OF INVENTION:  Methods of Inhibiting Angiogenesis and
;    TITLE OF INVENTION:  Ameliorating Cancer By Using Superfibronectin
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell & Flores LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/717,169
;      FILING DATE:  20-SEP-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-LJ 2017
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-717-169-17


  Query Match             100.0%;  Score 65;  DB 2;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.004;
```

```
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CDCRGDCFC 9
          |||||||||
Db      2 CDCRGDCFC 10


RESULT  13
US-08-286-861-10
; Sequence 10, Application US/08286861
; Patent No. 5981478
;  GENERAL INFORMATION:
;    APPLICANT:  Ruoslahti, Erkki
;    APPLICANT:  Koivunen, Erkki
;    TITLE OF INVENTION:  No. 5981478el Integrin-Binding Peptides
;    NUMBER OF SEQUENCES:  46
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/286,861
;      FILING DATE:  04-AUG-1994
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/158,001
;      FILING DATE:  24-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-LA 9992
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  circular
US-08-286-861-10


  Query Match           100.0%;  Score 65;  DB 2;  Length 11;
  Best Local Similarity  100.0%;  Pred. No. 0.004;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CDCRGDCFC 9
          |||||||||
Db      2 CDCRGDCFC 10


RESULT  14
US-09-139-802-16
; Sequence 16, Application US/09139802
; Patent No. 6180084
; GENERAL INFORMATION:
;  APPLICANT: Ruoslahti, Erkki
;  APPLICANT: Pasqualini, Renata
;  TITLE OF INVENTION: NGR Receptor and Methods of Identifying Tumor Homing
;  TITLE OF INVENTION: Molecules That Home to Angiogenic Vasculature Using
;  TITLE OF INVENTION: Same
;  FILE REFERENCE: P-LJ 3203
;  CURRENT APPLICATION NUMBER: US/09/139,802
;  CURRENT FILING DATE: 1998-08-25
;  EARLIER APPLICATION NUMBER: 08/926,914
;  EARLIER FILING DATE: 1997-09-10
;  EARLIER APPLICATION NUMBER: 08/710,067
;  EARLIER FILING DATE: 1996-09-10
;  NUMBER OF SEQ ID NOS: 226
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
;   LENGTH: 11
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: Peptide
US-09-139-802-16


  Query Match           100.0%;  Score 65;  DB 4;  Length 11;
  Best Local Similarity  100.0%;  Pred. No. 0.004;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CDCRGDCFC 9
          |||||||||
Db      2 CDCRGDCFC 10


RESULT  15
US-08-701-124-79
; Sequence 79, Application US/08701124
; Patent No. 5846782
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Roelvink, Petrus W.
;    APPLICANT:  Kovesdi, Imre
;    TITLE OF INVENTION:  TARGETING ADENOVIRUS WITH USE OF
;    TITLE OF INVENTION:  CONSTRAINED PEPTIDE MOTIFS
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza - 49th Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/701,124
;      FILING DATE:  21-AUG-1996
;  INFORMATION FOR SEQ ID NO:  79:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  12 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-701-124-79


  Query Match           100.0%;  Score 65;  DB 2;  Length 12;
  Best Local Similarity  100.0%;  Pred. No. 0.0043;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CDCRGDCFC 9
          |||||||||
Db      3 CDCRGDCFC 11


RESULT  16
US-09-130-225-79
; Sequence 79, Application US/09130225
; Patent No. 6057155
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Roelvink, Petrus W.
;    APPLICANT:  Kovesdi, Imre
;    TITLE OF INVENTION:  TARGETING ADENOVIRUS WITH USE OF
;    TITLE OF INVENTION:  CONSTRAINED PEPTIDE MOTIFS
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza - 49th Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/130,225
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 8-701124
;      FILING DATE:  21-AUG-1996
;  INFORMATION FOR SEQ ID NO:  79:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  12 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-130-225-79


  Query Match             100.0%;  Score 65;  DB 3;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 0.0043;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           |||||||||
Db       3 CDCRGDCFC 11


RESULT  17
US-09-455-061-79
; Sequence 79, Application US/09455061
; Patent No. 6329190
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Roelvink, Petrus W.
;    APPLICANT:  Kovesdi, Imre
;    TITLE OF INVENTION:  TARGETING ADENOVIRUS WITH USE OF
;    TITLE OF INVENTION:  CONSTRAINED PEPTIDE MOTIFS
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza - 49th Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/455,061
;      FILING DATE:  06-DEC-1999
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 9-130225
;      FILING DATE:  06-AUG-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 8-701124
;      FILING DATE:  21-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hefner, M. Daniel
;      REGISTRATION NUMBER:  41,826
;      REFERENCE/DOCKET NUMBER:  203128
;  INFORMATION FOR SEQ ID NO:  79:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  12 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-455-061-79


  Query Match             100.0%;  Score 65;  DB 4;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 0.0043;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           |||||||||
Db       3 CDCRGDCFC 11


RESULT  18
US-08-701-124-68
; Sequence 68, Application US/08701124
; Patent No. 5846782
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Roelvink, Petrus W.
;    APPLICANT:  Kovesdi, Imre
;    TITLE OF INVENTION:  TARGETING ADENOVIRUS WITH USE OF
;    TITLE OF INVENTION:  CONSTRAINED PEPTIDE MOTIFS
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza - 49th Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/701,124
;      FILING DATE:  21-AUG-1996
;  INFORMATION FOR SEQ ID NO:  68:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-701-124-68


  Query Match             100.0%;  Score 65;  DB 2;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 0.0049;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           |||||||||
Db       3 CDCRGDCFC 11


RESULT  19
US-09-130-225-68
```

```
; Sequence 68, Application US/09130225
; Patent No. 6057155
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Roelvink, Petrus W.
;    APPLICANT:  Kovesdi, Imre
;    TITLE OF INVENTION:  TARGETING ADENOVIRUS WITH USE OF
;    TITLE OF INVENTION:  CONSTRAINED PEPTIDE MOTIFS
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza - 49th Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/130,225
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 8-701124
;      FILING DATE:  21-AUG-1996
;  INFORMATION FOR SEQ ID NO:  68:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-130-225-68


  Query Match           100.0%;  Score 65;  DB 3;  Length 14;
  Best Local Similarity  100.0%;  Pred. No. 0.0049;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           |||||||||
Db       3 CDCRGDCFC 11


RESULT  20
US-09-455-061-68
; Sequence 68, Application US/09455061
; Patent No. 6329190
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Roelvink, Petrus W.
;    APPLICANT:  Kovesdi, Imre
;    TITLE OF INVENTION:  TARGETING ADENOVIRUS WITH USE OF
;    TITLE OF INVENTION:  CONSTRAINED PEPTIDE MOTIFS
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza - 49th Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/455,061
;      FILING DATE:  06-DEC-1999
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 9-130225
;      FILING DATE:  06-AUG-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 8-701124
;      FILING DATE:  21-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hefner, M. Daniel
;      REGISTRATION NUMBER:  41,826
;      REFERENCE/DOCKET NUMBER:  203128
;  INFORMATION FOR SEQ ID NO:  68:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-455-061-68


  Query Match           100.0%;  Score 65;  DB 4;  Length 14;
  Best Local Similarity  100.0%;  Pred. No. 0.0049;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           |||||||||
Db       3 CDCRGDCFC 11


RESULT  21
US-08-701-124-31
; Sequence 31, Application US/08701124
; Patent No. 5846782
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Roelvink, Petrus W.
;    APPLICANT:  Kovesdi, Imre
;    TITLE OF INVENTION:  TARGETING ADENOVIRUS WITH USE OF
;    TITLE OF INVENTION:  CONSTRAINED PEPTIDE MOTIFS
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza - 49th Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/701,124
;      FILING DATE:  21-AUG-1996
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-701-124-31


  Query Match           100.0%;  Score 65;  DB 2;  Length 15;
  Best Local Similarity  100.0%;  Pred. No. 0.0052;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           |||||||||
Db       4 CDCRGDCFC 12


RESULT  22
```

```
US-09-130-225-31
; Sequence 31, Application US/09130225
; Patent No. 6057155
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Roelvink, Petrus W.
;    APPLICANT:  Kovesdi, Imre
;    TITLE OF INVENTION:  TARGETING ADENOVIRUS WITH USE OF
;    TITLE OF INVENTION:  CONSTRAINED PEPTIDE MOTIFS
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza - 49th Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/130,225
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 8-701124
;      FILING DATE:  21-AUG-1996
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-130-225-31


  Query Match           100.0%;  Score 65;  DB 3;  Length 15;
  Best Local Similarity  100.0%;  Pred. No. 0.0052;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           |||||||||
Db       4 CDCRGDCFC 12


RESULT  23
US-09-426-680-7
;Sequence 7, Application US/09426680
;Patent No. 6287857
; GENERAL INFORMATION:
;  APPLICANT: Catherine R. O'Riordan
;  APPLICANT: Samuel C. Wadsworth
;  TITLE OF INVENTION: Nucleic Acid Delivery Vehicles
;  FILE REFERENCE: GA0103USB2
;  CURRENT APPLICATION NUMBER: US/09/426,680
;  CURRENT FILING DATE: 1999-10-25
;  EARLIER APPLICATION NUMBER: PCT/US99/02680
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: human
;   FEATURE:
;   NAME/KEY: DISULFID
;   LOCATION: (0)...(0)
;   NAME/KEY: PEPTIDE
;   LOCATION: (0)...(0)
US-09-426-680-7


  Query Match           100.0%;  Score 65;  DB 4;  Length 15;
  Best Local Similarity  100.0%;  Pred. No. 0.0052;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           |||||||||
Db       3 CDCRGDCFC 11


RESULT  24
US-09-455-061-31
; Sequence 31, Application US/09455061
; Patent No. 6329190
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Roelvink, Petrus W.
;    APPLICANT:  Kovesdi, Imre
;    TITLE OF INVENTION:  TARGETING ADENOVIRUS WITH USE OF
;    TITLE OF INVENTION:  CONSTRAINED PEPTIDE MOTIFS
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza - 49th Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/455,061
;      FILING DATE:  06-DEC-1999
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 9-130225
;      FILING DATE:  06-AUG-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 8-701124
;      FILING DATE:  21-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hefner, M. Daniel
;      REGISTRATION NUMBER:  41,826
;      REFERENCE/DOCKET NUMBER:  203128
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-455-061-31


  Query Match           100.0%;  Score 65;  DB 4;  Length 15;
  Best Local Similarity  100.0%;  Pred. No. 0.0052;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CDCRGDCFC 9
           |||||||||
Db       4 CDCRGDCFC 12


RESULT  25
US-08-701-124-49
; Sequence 49, Application US/08701124
; Patent No. 5846782
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Roelvink, Petrus W.
;    APPLICANT:  Kovesdi, Imre
;    TITLE OF INVENTION:  TARGETING ADENOVIRUS WITH USE OF
;    TITLE OF INVENTION:  CONSTRAINED PEPTIDE MOTIFS
;    NUMBER OF SEQUENCES:  80
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza - 49th Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/701,124
;      FILING DATE:  21-AUG-1996
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-701-124-49


  Query Match           100.0%;  Score 65;  DB 2;  Length 24;
  Best Local Similarity  100.0%;  Pred. No. 0.0077;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CDCRGDCFC 9
          |||||||||
Db     15 CDCRGDCFC 23


RESULT  26
US-09-130-225-49
; Sequence 49, Application US/09130225
; Patent No. 6057155
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Roelvink, Petrus W.
;    APPLICANT:  Kovesdi, Imre
;    TITLE OF INVENTION:  TARGETING ADENOVIRUS WITH USE OF
;    TITLE OF INVENTION:  CONSTRAINED PEPTIDE MOTIFS
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza - 49th Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/130,225
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 8-701124
;      FILING DATE:  21-AUG-1996
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-130-225-49


  Query Match           100.0%;  Score 65;  DB 3;  Length 24;
  Best Local Similarity  100.0%;  Pred. No. 0.0077;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CDCRGDCFC 9
          |||||||||
Db     15 CDCRGDCFC 23


RESULT  27
US-09-455-061-49
; Sequence 49, Application US/09455061
; Patent No. 6329190
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Roelvink, Petrus W.
;    APPLICANT:  Kovesdi, Imre
;    TITLE OF INVENTION:  TARGETING ADENOVIRUS WITH USE OF
;    TITLE OF INVENTION:  CONSTRAINED PEPTIDE MOTIFS
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza - 49th Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/455,061
;      FILING DATE:  06-DEC-1999
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 9-130225
;      FILING DATE:  06-AUG-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 8-701124
;      FILING DATE:  21-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hefner, M. Daniel
;      REGISTRATION NUMBER:  41,826
;      REFERENCE/DOCKET NUMBER:  203128
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-455-061-49


  Query Match           100.0%;  Score 65;  DB 4;  Length 24;
  Best Local Similarity  100.0%;  Pred. No. 0.0077;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CDCRGDCFC 9
          |||||||||
Db     15 CDCRGDCFC 23


Search completed: May 29, 2002, 10:43:04
Job time: 22 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 29, 2002, 09:55:26 ; Search time 19.72 Seconds
(without alignments)
43.854 Million cell updates/sec

Title: US-09-734-628-1
Perfect score: 65
Sequence: 1 CDCRGDCFC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 788

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	25	38.5	8	2	S59622	metallothionein is
2	23	35.4	5	2	B45525	actin I - malaria
3	21	32.3	6	2	PT0652	T-cell receptor be
4	20	30.8	5	2	A33882	cadmium-binding pe
5	20	30.8	7	2	B33882	cadmium-binding he
6	17	26.2	5	2	F22565	R-phycoerythrin ga
7	17	26.2	6	2	I67345	MHC H2-K-k cell su
8	17	26.2	9	2	PH0942	T-cell receptor be
9	16.5	25.4	7	2	A12016	formylglycinamide
10	16	24.6	7	2	A58512	venom heptapeptide
11	16	24.6	7	2	PT0620	T-cell receptor be
12	16	24.6	8	2	PC1002	leucine--tRNA liga
13	15	23.1	6	2	I37263	Y protein - human
14	15	23.1	7	2	PH1408	Ig heavy chain V r
15	15	23.1	7	2	S38516	mabinlin II chain
16	15	23.1	8	2	PH1407	Ig heavy chain V r
17	15	23.1	9	2	A28495	conopressin G - co
18	15	23.1	9	2	S19329	sperm-activating p
19	15	23.1	9	2	PC2021	oxytocin-related p
20	15	23.1	9	2	A26363	cardioactive pepti
21	15	23.1	9	2	S27233	cardioactive pepti
22	15	23.1	9	2	S39767	cardioactive pepti
23	15	23.1	9	2	S39040	lysine-conopressin
24	14	21.5	4	2	S43959	Ig mu chain V regi
25	14	21.5	6	2	A41946	T-cell receptor ga
26	14	21.5	8	2	A37521	R-phycoerythrin ga
27	14	21.5	8	2	S11078	glucose-6-phosphat
28	14	21.5	8	2	PT0279	Ig heavy chain CRD
29	14	21.5	8	2	I57018	gene Cftr protein
30	14	21.5	9	2	B46250	alpha-adaptin - bo
31	13	20.0	6	2	JU0355	lipopeptide WS1279
32	13	20.0	7	4	I55382	hypothetical pepti
33	13	20.0	9	2	A60522	sperm-activating p
34	13	20.0	9	2	QDRB	delta sleep-induci
35	13	20.0	9	2	A12872	transaldolase (EC
36	13	20.0	9	2	PT0634	T-cell receptor be
37	12	18.5	3	3	A22565	R-phycoerythrin al
38	12	18.5	4	2	PT0711	T-cell receptor be
39	12	18.5	5	2	PQ0689	photosystem I 10.4
40	12	18.5	5	2	PT0513	T-cell receptor be
41	12	18.5	5	2	PT0538	T-cell receptor be
42	12	18.5	5	2	PT0703	T-cell receptor be
43	12	18.5	5	2	PT0690	T-cell receptor be
44	12	18.5	5	2	PT0573	T-cell receptor be
45	12	18.5	5	2	PT0679	T-cell receptor be

ALIGNMENTS

Closed

:admium-binding - Arianta arbustorum (terrestrial snail) (
m
e_revision 24-Oct-1997 #text_change 07-May-1999

; Hauer, C.R.; Birchler, N.; Dallinger, R.
95
id amino acid sequencing of two cadmium-binding metalloth
fUID:96067616

n
.. ...ciation; metal binding; metal-thiolate cluster

Query Match 38.5%; Score 25; DB 2; Length 8;
Best Local Similarity 42.9%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 CRGDCFC 9
| | | |
Db 1 CNSSCSC 7

RESULT 2
B45525
actin I - malaria parasite (Plasmodium falciparum) (fragments)
C;Species: Plasmodium falciparum
C;Date: 03-Jun-1993 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
C;Accession: B45525
R;Wesseling, J.G.; Snijders, P.J.F.; van Someren, P.; Jansen, J.; Smits, M.A.; Schoen
Mol. Biochem. Parasitol. 35, 167-176, 1989
A;Title: Stage-specific expression and genomic organization of the actin genes of the
A;Reference number: A45525; MUID:89364996
A;Accession: B45525
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-5 <WES>
A;Cross-references: GB:J03988
A;Note: the authors translated the codon GAA for residue 3 as Gly
C;Comment: The actin I gene contains no introns.

Query Match 35.4%; Score 23; DB 2; Length 5;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GDCF 8
|:||
Db 2 GECF 5

```
RESULT   3
PT0652
T-cell receptor beta chain V-D-J region (121-1E) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0652
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0652
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FEE>
A;Experimental source: day 4 postnatal thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             32.3%;  Score 21;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 GDC 7
           |||
Db       3 GDC 5


RESULT   4
A33882
cadmium-binding pentapeptide - downy thornapple
C;Species: Datura innoxia (downy thornapple)
C;Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 18-Jun-1993
C;Accession: A33882
R;Jackson, P.J.; Unkefer, C.J.; Doolen, J.A.; Watt, K.; Robinson, N.J.
Proc. Natl. Acad. Sci. U.S.A. 84, 6619-6623, 1987
A;Title: Poly(gamma-glutamylcysteinyl)glycine: its role in cadmium resistance in plant c
A;Reference number: A94182; MUID:88016144
A;Accession: A33882
A;Molecule type: protein
A;Residues: 1-5 <JAC>

  Query Match             30.8%;  Score 20;  DB 2;  Length 5;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CDC 3
           |:|
Db       2 CEC 4


RESULT   5
B33882
cadmium-binding heptapeptide - downy thornapple
C;Species: Datura innoxia (downy thornapple)
C;Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 18-Jun-1993
C;Accession: B33882
R;Jackson, P.J.; Unkefer, C.J.; Doolen, J.A.; Watt, K.; Robinson, N.J.
Proc. Natl. Acad. Sci. U.S.A. 84, 6619-6623, 1987
A;Title: Poly(gamma-glutamylcysteinyl)glycine: its role in cadmium resistance in plant c
A;Reference number: A94182; MUID:88016144
A;Accession: B33882
A;Molecule type: protein
A;Residues: 1-7 <JA2>

  Query Match             30.8%;  Score 20;  DB 2;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CDC 3
           |:|
Db       2 CEC 4


RESULT   6
F22565
R-phycoerythrin gamma-A chain - red alga (Gastroclonium coulteri) (fragment)
C;Species: Gastroclonium coulteri
C;Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 23-Mar-1993
C;Accession: F22565
R;Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A;Title: Characterization of the bilin attachment sites in R-phycoerythrin.
A;Reference number: A22565; MUID:85182601
A;Accession: F22565
A;Molecule type: protein
A;Residues: 1-5 <KLO>

  Query Match             26.2%;  Score 17;  DB 2;  Length 5;
  Best Local Similarity   50.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 GDCF 8
           | |:
Db       1 GTCY 4


RESULT   7
I67345
MHC H2-K-k cell surface glycoprotein - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
C;Accession: I67345
R;Archibald, A.L.; Thompson, N.A.; Kvist, S.
EMBO J. 5, 957-965, 1986
A;Title: A single nucleotide difference at the 3' end of an intron causes differentia
A;Reference number: I53243; MUID:86247587
A;Accession: I67345
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-6 <RES>
A;Cross-references: GB:M26859; NID:g199439; PIDN:AAA39612.1; PID:g387458
C;Genetics:
A;Introns: 6/1
C;Keywords: glycoprotein

  Query Match             26.2%;  Score 17;  DB 2;  Length 6;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 DCR 4
           ||:
Db       3 DCK 5


RESULT   8
PH0942
T-cell receptor beta chain V-D-J region (clone 13) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 30-May-1997
C;Accession: PH0942
R;Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.
J. Exp. Med. 174, 1467-1476, 1991
A;Title: Analysis of T cell receptor beta chains in Lewis rats with experimental alle
A;Reference number: PH0891; MUID:92078857
A;Accession: PH0942
A;Molecule type: mRNA
A;Residues: 1-9 <GOL>
A;Experimental source: complete Freund's adjuvant-immunized lymph node
A;Note: the authors translated the codon TGC for residue 2 as Ala
C;Keywords: T-cell receptor
```

Query Match 26.2%; Score 17; DB 2; Length 9;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy      3 CRG 5
          |:|
Db      2 CKG 4
```

RESULT 9
A12016
formylglycinamide ribonucleotide amidotransferase (EC 2.-.-.-) - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 13-Mar-1997
C;Accession: A12016; B12016
R;Ohnoki, S.; Hong, B.S.; Buchanan, J.M.
Fed. Proc. 35, 1549, 1976
A;Title: Amino acid sequence at glutamine active site for FGAR-amidotransferase.
A;Reference number: A91459
A;Accession: A12016
A;Molecule type: protein
A;Residues: 1-7 <OHN>
A;Experimental source: liver, peptide 1
A;Accession: B12016
A;Molecule type: protein
A;Residues: 1-5 <OH2>
A;Experimental source: liver, peptide 2
C;Keywords: transferase

Query Match 25.4%; Score 16.5; DB 2; Length 7;
Best Local Similarity 42.9%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 3; Gaps 1;

```
Qy      1 CDCRGDC 7
          ||   :|
Db      3 CD---BC 6
```

RESULT 10
A58512
venom heptapeptide - cone shell (Conus imperialis)
C;Species: Conus imperialis (imperial cone)
C;Date: 19-Mar-1997 #sequence_revision 11-Apr-1997 #text_change 07-May-1999
C;Accession: A58512
R;Craig, A.G.; Jimenez, E.C.; Dykert, J.; Nielsen, D.B.; Gulyas, J.; Abogadie, F.C.; Por
J. Biol. Chem. 272, 4689-4698, 1997
A;Title: A novel post-translational modification involving bromination of tryptophan. Ic
A;Reference number: A58512; MUID:97184108
A;Accession: A58512
A;Molecule type: protein
A;Residues: 1-7 <CRA>
C;Superfamily: unassigned conotoxins
C;Keywords: amidated carboxyl end; bromine; pyroglutamic acid; venom
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;6/Modified site: 6-bromotryptophan (Trp) #status experimental
F;7/Modified site: amidated carboxyl end (Cys) #status experimental

Query Match 24.6%; Score 16; DB 2; Length 7;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy      5 GDCFC 9
          |  :|
Db      3 GQAWC 7
```

RESULT 11
PT0620
T-cell receptor beta chain V-D-J region (120-2CO) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0620
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions
A;Reference number: PT0509; MUID:91277601
A;Accession: PT0620
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-7 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 24.6%; Score 16; DB 2; Length 7;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      2 DCRG 5
          | ||
Db      4 DVRG 7
```

RESULT 12
PC1002
leucine--tRNA ligase (EC 6.1.1.4) - Escherichia coli (fragments)
N;Alternate names: leucyl-tRNA synthetase
C;Species: Escherichia coli
C;Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 26-Feb-1998
C;Accession: PC1002
R;Miao, F.; Shi, J.P.; Wang, Y.L.
Science in China (series B) 34, 691-698, 1991
A;Title: Chemical modification of sulfhydryl groups of E. coli leucyl-tRNA synthetase
A;Reference number: PC1002
A;Accession: PC1002
A;Molecule type: protein
A;Residues: 1-8 <MIA>
C;Comment: This enzyme catalyzes the aminoacylation of tRNA(leu) with Leucine.
C;Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis
F;5-8/Region: catalytic #status predicted

Query Match 24.6%; Score 16; DB 2; Length 8;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 CDCR 4
          || :
Db      5 CDTK 8
```

RESULT 13
I37263
Y protein - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 05-Nov-1999
C;Accession: I37263
R;Waeber, G.; Habener, J.F.
Endocrinology 131, 2010-2015, 1992
A;Title: Novel testis germ cell-specific transcript of the CREB gene contains an alte
A;Reference number: I37263; MUID:93010691
A;Accession: I37263
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-6 <RES>
A;Cross-references: EMBL:X68994; NID:g396171; PIDN:CAA48780.1; PID:g579816
C;Genetics:
A;Gene: CREB

Query Match 23.1%; Score 15; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       8 FC 9
           ||
Db       3 FC 4
```

RESULT 14
PH1408
Ig heavy chain V region - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Mar-1999
C;Accession: PH1408; PH1405
R;Shirasawa, T.; Miyazoe, I.; Hagiwara, S.; Kimoto, H.; Shigemoto, K.; Taniguchi, M.; Ta
J. Exp. Med. 176, 1209-1214, 1992
A;Title: Heavy chain variable (VH) region diversity generated by VH gene replacement in
ia virus.
A;Reference number: PH1403; MUID:93018837
A;Accession: PH1408
A;Molecule type: DNA
A;Residues: 1-7 <SHI>
A;Experimental source: clone micro m+ 46-12-2
A;Accession: PH1405
A;Molecule type: DNA
A;Residues: 1-7 <SHI2>
A;Experimental source: clone micro m+ 46-6
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

Query Match 23.1%; Score 15; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       8 FC 9
           ||
Db       1 FC 2
```

RESULT 15
S38516
mabinlin II chain A - Yunnan caper (fragments)
C;Species: Capparis masaikai (Yunnan caper)
C;Date: 08-Jun-1994 #sequence_revision 27-Feb-1997 #text_change 13-Mar-1997
C;Accession: S38516
R;Nirasawa, S.; Liu, X.; Nishino, T.; Kurihara, Y.
Biochim. Biophys. Acta 1202, 277-280, 1993
A;Title: Disulfide bridge structure of the heat-stable sweet protein mabinlin II.
A;Reference number: S38516; MUID:94002261
A;Accession: S38516
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <NIR>

Query Match 23.1%; Score 15; DB 2; Length 7;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy       3 CRGDC 7
           |:  |
Db       1 CQRAC 5
```

Search completed: May 29, 2002, 09:58:09
Job time: 163 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 29, 2002, 09:57:41 ; Search time 9.77 Seconds
(without alignments)
35.668 Million cell updates/sec

Title: US-09-734-628-1
Perfect score: 65
Sequence: 1 CDCRGDCFC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 231

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	21	32.3	9	1	COVN_CONVE	P83047 conus ventr
2	15	23.1	7	1	FAR1_HELTI	P41871 helisoma tr
3	15	23.1	8	1	ACT_CARMA	P80709 carcinus ma
4	15	23.1	9	1	CCAP_CARMA	P38556 carcinus ma
5	15	23.1	9	1	CONO_CONGE	P05486 conus geogr
6	15	23.1	9	1	OXYT_EISFO	P42998 eisenia foe
7	15	23.1	9	1	SAP_STOVA	P24047 stomopneute
8	13	20.0	9	1	DSIP_RABIT	P01158 oryctolagus
9	13	20.0	9	1	TAL1_PICJA	P17440 pichia jadi
10	12	18.5	4	1	OCP1_OCTMI	P58648 octopus min
11	12	18.5	5	1	UXA4_CHLTR	P38005 chlamydia t
12	12	18.5	8	1	ACI_THUAL	P18691 thunnus alb
13	12	18.5	8	1	LMT2_LOCMI	P22396 locusta mig
14	12	18.5	8	1	ORMY_ORCLI	P82455 orconectes
15	12	18.5	9	1	ISOT_CYPCA	P42993 cyprinus ca
16	12	18.5	9	1	OXYA_SCYCA	P42996 scyliorhinu
17	12	18.5	9	1	OXYA_SQUAC	P42999 squalus aca
18	12	18.5	9	1	OXYF_SCYCA	P42997 scyliorhinu
19	12	18.5	9	1	OXYT_BUFRE	P42995 bufo regula
20	12	18.5	9	1	OXYT_CYPCA	P23879 cyprinus ca
21	12	18.5	9	1	OXYT_OCTVU	P80027 octopus vul
22	12	18.5	9	1	OXYT_RABIT	P32878 oryctolagus
23	12	18.5	9	1	OXYT_RAJCL	P42994 raja clavat
24	12	18.5	9	1	OXYV_SQUAC	P43000 squalus aca
25	12	18.5	9	1	UPA6_HUMAN	P30092 homo sapien
26	11	16.9	6	1	FARP_MONEX	P41966 moniezia ex
27	11	16.9	7	1	UF04_MOUSE	P38642 mus musculu
28	11	16.9	8	1	ALL5_CYDPO	P82156 cydia pomon
29	11	16.9	8	1	GLUR_HUMAN	P02729 homo sapien
30	11	16.9	9	1	CONO_CONST	P05487 conus stria
31	11	16.9	9	1	DNF1_LOCMI	P16339 locusta mig
32	11	16.9	9	1	FIBB_PAPAN	P19344 papio anubi
33	11	16.9	9	1	PGLR_DIAAB	P81179 diaprepes a
34	11	16.9	9	1	ULAH_HUMAN	P31934 homo sapien
35	10	15.4	5	1	RE11_LITRU	P82070 litoria rub
36	10	15.4	8	1	RS1_ERWCH	P37985 erwinia chr
37	10	15.4	9	1	TAL3_PICJA	P17441 pichia jadi
38	9	13.8	7	1	FARB_CALVO	P41866 calliphora
39	9	13.8	8	1	CCKN_MACEU	P30369 macropus eu
40	9	13.8	8	1	LCK5_LEUMA	P19987 leucophaea
41	9	13.8	8	1	UC26_MAIZE	P80632 zea mays (m
42	9	13.8	9	1	D1_NEPNO	P24816 nephrops no
43	9	13.8	9	1	FAR5_CALVO	P41860 calliphora
44	9	13.8	9	1	FAR6_CALVO	P41861 calliphora
45	9	13.8	9	1	FAR7_CALVO	P41862 calliphora

ALIGNMENTS

RESULT 1
COVN_CONVE
ID COVN_CONVE STANDARD; PRT; 9 AA.
AC P83047;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Contryphan-Vn.
OS Conus ventricosus (Mediterranean cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=117992;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RA Raybaudi Massilia G., Schinina M.E., Ascenzi P., Polticelli F.;
RT "Contryphan-Vn: a novel peptide from the venom of the Mediterranean
RT vermivorous marine snail Conus ventricosus.";
RL Submitted (JUL-2001) to the SWISS-PROT data bank.
CC -!- TISSUE SPECIFICITY: Venom.
KW Amidation; Venom; D-amino acid.
FT DISULFID 3 9
FT MOD_RES 5 5 D-TRYPTOPHAN.
FT MOD_RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 1091 MW; 8D38676323676EBA CRC64;

Query Match 32.3%; Score 21; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       5 GDC 7
           |||
Db       1 GDC 3
```

RESULT 2
FAR1_HELTI
ID FAR1_HELTI STANDARD; PRT; 7 AA.
AC P41871;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FMRFamide-like neuropeptide GDPFLRF-amide.
OS Helisoma trivolvis (Snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC Planorbidae; Helisoma.
OX NCBI_TaxID=27815;
RN [1]
RP SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=94286417; PubMed=7912428;
RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
RT "FMRFamide-related peptides from the kidney of the snail, Helisoma
RT trivolvis.";
RL Peptides 15:31-36(1994).

```
CC   -!- FUNCTION: APPEARS TO BE INVOLVED IN OSMOREGULATION BY AFFECTING
CC       THE KIDNEY, MANTLE AND SKIN.
CC   -!- TISSUE SPECIFICITY: KIDNEY, SKIN, MANTLE AND THE HEMOLYMPH.
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       7      7       AMIDATION.
SQ   SEQUENCE   7 AA;  851 MW;  69D40729D76AA810 CRC64;

  Query Match             23.1%;  Score 15;  DB 1;  Length 7;
  Best Local Similarity   75.0%;  Pred. No. 1e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 GDCF 8
          || |
Db      1 GDPF 4


RESULT   3
ACT_CARMA
ID   ACT_CARMA      STANDARD;      PRT;     8 AA.
AC   P80709;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Actin (Fragment).
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RA   Lachaise F., Somme G., Carpentier G., Granjeon E., Webster S.,
RA   Baghdassarian D.;
RT   "A transaldolase. An enzyme implicated in crab steroidogenesis.";
RL   Endocrine 5:23-32(1996).
CC   -!- FUNCTION: ACTINS ARE HIGHLY CONSERVED PROTEINS THAT ARE INVOLVED
CC       IN VARIOUS TYPES OF CELL MOTILITY AND ARE UBIQUITOUSLY EXPRESSED
CC       IN ALL EUKARYOTIC CELLS.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:
CC       6.8, ITS MW IS: 46 kDa.
CC   -!- SIMILARITY: BELONGS TO THE ACTIN FAMILY.
DR   InterPro; IPR004001; Actin.
DR   InterPro; IPR004000; Actin_like.
DR   PROSITE; PS00406; ACTINS_1; PARTIAL.
DR   PROSITE; PS00432; ACTINS_2; PARTIAL.
DR   PROSITE; PS01132; ACTINS_ACT_LIKE; PARTIAL.
KW   Structural protein.
FT   NON_TER       1      1
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  976 MW;  1424005AB2CAAEB3 CRC64;

  Query Match             23.1%;  Score 15;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CD 2
          ||
Db      2 CD 3


RESULT   4
CCAP_CARMA
ID   CCAP_CARMA     STANDARD;      PRT;     9 AA.
AC   P38556;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Cardioactive peptide (CCAP).
OS   Carcinus maenas (Common shore crab) (Green crab),
OS   Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm),
OS   Tenebrio molitor (Yellow mealworm), and
OS   Spodoptera eridania (Southern armyworm).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759, 7130, 7067, 37547;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=C.maenas; TISSUE=Pericardial organs;
RA   Stangier J., Hilbich C., Beyreuther K., Keller R.;
RT   "Unusual cardioactive peptide (CCAP) from pericardial organs of the
RT   shore crab Carcinus maenas.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:575-579(1987).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=M.sexta;
RX   MEDLINE=93050243; PubMed=1426284;
RA   Cheung C.C., Loi P.K., Sylwester A.W., Lee T.D., Tublitz N.J.;
RT   "Primary structure of a cardioactive neuropeptide from the tobacco
RT   hawkmoth, Manduca sexta.";
RL   FEBS Lett. 313:165-168(1992).
RN   [3]
RP   SEQUENCE.
RC   SPECIES=T.molitor, and S.eridania; TISSUE=Head;
RX   MEDLINE=94176032; PubMed=8129851;
RA   Furuya K., Liao S., Reynolds S.E., Ota R.B., Hackett M.,
RA   Schooley D.A.;
RT   "Isolation and identification of a cardioactive peptide from Tenebrio
RT   molitor and Spodoptera eridania.";
RL   Biol. Chem. Hoppe-Seyler 374:1065-1074(1993).
CC   -!- FUNCTION: THE EFFECT OF CCAP IS BOTH INO- AND CHRONOTROPIC.
CC   -!- TISSUE SPECIFICITY: STORED IN PERICARDIAL ORGANS AND RELEASED
CC       INTO THE HEMOLYMPH.
DR   PIR; A26363; A26363.
DR   PIR; S27233; S27233.
KW   Neuropeptide; Amidation.
FT   DISULFID      3      9
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  959 MW;  C5A861A9CDD44EB9 CRC64;

  Query Match             23.1%;  Score 15;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      8 FC 9
          ||
Db      2 FC 3


RESULT   5
CONO_CONGE
ID   CONO_CONGE     STANDARD;      PRT;     9 AA.
AC   P05486;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   LYS-conopressin G.
OS   Conus geographus (Geography cone).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC   Neogastropoda; Conoidea; Conidae; Conus.
OX   NCBI_TaxID=6491;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=88058932; PubMed=3680228;
RA   Cruz L.J., de Santos V., Zafaralla G.C., Ramilo C.A., Zeikus R.D.,
RA   Gray W.R., Olivera B.M.;
RT   "Invertebrate vasopressin/oxytocin homologs. Characterization of
RT   peptides from Conus geographus and Conus straitus venoms.";
RL   J. Biol. Chem. 262:15821-15824(1987).
RN   [2]
```

```
RP   REVIEW.
RX   MEDLINE=89024586; PubMed=3052286;
RA   Gray W.R., Olivera B.M., Cruz L.J.;
RT   "Peptide toxins from venomous Conus snails.";
RL   Annu. Rev. Biochem. 57:665-700(1988).
CC   -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR   PIR; A28495; A28495.
DR   InterPro; IPR000981; Neurhypophys_horm.
DR   Pfam; PF00220; hormone4; 1.
DR   PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW   Hormone; Amidation.
FT   DISULFID       1      6
FT   MOD_RES        9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1037 MW;  D4FC276EB4540059 CRC64;


  Query Match             23.1%;  Score 15;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 CF 8
          ||
Db      1 CF 2


RESULT   6
OXYT_EISFO
ID   OXYT_EISFO     STANDARD;      PRT;     9 AA.
AC   P42998;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Annetocin.
OS   Eisenia foetida (Common brandling worm) (Common dung-worm).
OC   Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
OC   Lumbricina; Lumbricidae; Eisenia.
OX   NCBI_TaxID=6396;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Pituitary;
RX   MEDLINE=94121660; PubMed=8292046;
RA   Oumi T., Ukena K., Matsushima O., Ikeda T., Fujita T., Minakata H.,
RA   Nomoto K.;
RT   "Annetocin: an oxytocin-related peptide isolated from the earthworm,
RT   Eisenia foetida.";
RL   Biochem. Biophys. Res. Commun. 198:393-399(1994).
CC   -!- FUNCTION: POTENTIATES SPONTANEOUS CONTRACTIONS OF THE GUT AND ALSO
CC       PULSATORY CONTRACTIONS AND BLADDER-SHAKING MOVEMENT OF THE
CC       NEPHRIDIA. MAY BE INVOLVED IN OSMOREGULATION OF THE ANIMAL THROUGH
CC       NEPHRIDIAL FUNCTION.
CC   -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR   PIR; PC2021; PC2021.
DR   InterPro; IPR000981; Neurhypophys_horm.
DR   Pfam; PF00220; hormone4; 1.
DR   PROSITE; PS00264; NEUROHYPOPHYS_HORM; FALSE_NEG.
KW   Hormone; Amidation.
FT   DISULFID       1      6
FT   MOD_RES        9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  996 MW;  D4EEB76EB45412C9 CRC64;


  Query Match             23.1%;  Score 15;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 CF 8
          ||
Db      1 CF 2


RESULT   7
SAP_STOVA
ID   SAP_STOVA      STANDARD;      PRT;     9 AA.
AC   P24047;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   01-MAR-1992 (Rel. 21, Last annotation update)
DE   Sperm-activating peptide (SAP).
OS   Stomopneutes variolaris (Sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Diadematacea; Phymosomatoida; Stomechinidae;
OC   Stomopneustes.
OX   NCBI_TaxID=7663;
RN   [1]
RP   SEQUENCE, AND DISULFIDE BOND.
RC   TISSUE=Egg jelly;
RX   MEDLINE=92097763; PubMed=1756858;
RA   Yoshimo K.-I., Takao T., Shimonishi Y., Suzuki N.;
RT   "Determination of the amino acid sequence of an intramolecular
RT   disulfide linkage-containing sperm-activating peptide by tandem mass
RT   spectrometry.";
RL   FEBS Lett. 294:179-182(1991).
CC   -!- FUNCTION: CAUSE STIMULATION OF SPERM RESPIRATION AND MOTILITY
CC       THROUGH INTRACELLULAR ALKALINIZATION, TRANSIENT ELEVATIONS OF
CC       CAMP, CGMP AND CLACIUM LEVELS IN SPERM CELLS, AND TRANSIENT
CC       ACTIVATION AND SUBSEQUENT INACTIVATION OF THE MEMBRANE FORM OF
CC       GUANYLATE CYCLASE.
DR   PIR; S19329; S19329.
FT   DISULFID       3      8
SQ   SEQUENCE   9 AA;  1010 MW;  C469B3387B076EB9 CRC64;


  Query Match             23.1%;  Score 15;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      8 FC 9
          ||
Db      2 FC 3


RESULT   8
DSIP_RABIT
ID   DSIP_RABIT     STANDARD;      PRT;     9 AA.
AC   P01158;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Delta sleep-inducing peptide (DSIP).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=77185324; PubMed=862769;
RA   Monnier M., Dudler L., Gachter R., Maier P.F., Tobler H.J.,
RA   Schoenenberger G.A.;
RT   "The delta sleep inducing peptide (DSIP). Comparative properties of
RT   the original and synthetic nonapeptide.";
RL   Experientia 33:548-552(1977).
RN   [2]
RP   SEQUENCE, AND SYNTHESIS.
RX   MEDLINE=79054421; PubMed=568769;
RA   Schoenenberger G.A., Maier P.F., Tobler H.J., Wilson K., Monnier M.;
RT   "The delta EEG (sleep)-inducing peptide (DSIP). XI. Amino-acid
RT   analysis, sequence, synthesis and activity of the nonapeptide.";
RL   Pflugers Arch. 376:119-129(1978).
RN   [3]
RP   REVIEW.
RX   MEDLINE=87175129; PubMed=3550726;
RA   Graf M.V., Kastin A.J.;
RT   "Delta-sleep-inducing peptide (DSIP): an update.";
RL   Peptides 7:1165-1187(1986).
CC   -!- FUNCTION: WHEN INFUSED INTO THE MESODIENCEPHALIC VENTRICLE OF
CC       RECIPIENT RABBITS INDUCES SPINDLE AND DELTA EEG ACTIVITY AND
```

```
CC       REDUCED MOTOR ACTIVITIES.
CC   -!- MISCELLANEOUS: THIS PEPTIDE WAS OBTAINED FROM DIALYSATES OF
CC       OCCIPITAL VENOUS SINUS BLOOD FROM RABBITS KEPT ASLEEP BY ELECTRIC
CC       STIMULATION OF THE THALAMUS.
CC   -!- DATABASE: NAME=Protein Spotlight;
CC       NOTE=Issue 8 of March 2001;
CC       WWW="http://www.expasy.org/spotlight/articles/sptlt008.html".
DR   PIR; A01422; QDRB.
SQ   SEQUENCE   9 AA;  849 MW;  DDD365BDDAA8787D CRC64;


  Query Match             20.0%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   40.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 DCRGD 6
          |  |:
Db      5 DASGE 9


SULT   9
AL1_PICJA
ID   TAL1_PICJA     STANDARD;      PRT;     9 AA.
AC   P17440;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Transaldolase I (EC 2.2.1.2) (Fragment).
OS   Pichia jadinii (Yeast) (Candida utilis).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Pichia.
OX   NCBI_TaxID=4903;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=77110646; PubMed=556924;
RA   Sun S.C., Joris L., Tsolas O.;
RT   "Purification of crystallization of transaldolase isozyme I and
RT   evidence for different genetic origin of isozymes I and III in
RT   Candida utilis.";
RL   Arch. Biochem. Biophys. 178:69-78(1977).
CC   -!- FUNCTION: TRANSALDOLASE IS IMPORTANT FOR THE BALANCE OF
CC       METABOLITES IN THE PENTOSE-PHOSPHATE PATHWAY.
CC   -!- CATALYTIC ACTIVITY: Sedoheptulose 7-phosphate + D-glyceraldehyde
CC       3-phosphate = D-erythrose 4-phosphate + D-fructose 6-phosphate.
CC   -!- PATHWAY: Pentose phosphate pathway; nonoxidative part.
CC   -!- SIMILARITY: BELONGS TO THE TRANSALDOLASE FAMILY. SUBFAMILY 1.
DR   PIR; A12872; A12872.
DR   InterPro; IPR001585; Transaldolase.
DR   PROSITE; PS00958; TRANSALDOLASE_2; PARTIAL.
DR   PROSITE; PS01054; TRANSALDOLASE_1; PARTIAL.
KW   Transferase; Pentose shunt.
FT   NON_TER       1      1
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1008 MW;  274F31AF0EB1E058 CRC64;


  Query Match             20.0%;  Score 13;  DB 1;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1e+05;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CD 2
          |:
Db      5 CB 6


RESULT  10
OCP1_OCTMI
ID   OCP1_OCTMI     STANDARD;      PRT;     4 AA.
AC   P58648;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Cardioactive peptides Ocp-1/Ocp-2.
OS   Octopus minor (Octopus).
OC   Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;
OC   Incirrata; Octopodidae; Octopus.
OX   NCBI_TaxID=89766;
RN   [1]
RP   SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.
RC   TISSUE=Brain;
RX   PubMed=10876044;
RA   Iwakoshi E., Hisada M., Minakata H.;
RT   "Cardioactive peptides isolated from the brain of a Japanese octopus,
RT   Octopus minor.";
RL   Peptides 21:623-630(2000).
CC   -!- FUNCTION: Cardioactive; has both positive chronotropic and
CC       inotropic effects on the heart. Ocp-2 is a 1000 time less
CC       active than Ocp-1.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PTM: Ocp-2 has L-Phe instead of D-Phe.
CC   -!- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.
KW   Hormone; D-amino acid.
FT   MOD_RES       2      2       D-PHENYLALANINE.
SQ   SEQUENCE   4 AA;  394 MW;  6AA879C810000000 CRC64;


  Query Match             18.5%;  Score 12;  DB 1;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 GD 6
          ||
Db      3 GD 4


RESULT  11
UXA4_CHLTR
ID   UXA4_CHLTR     STANDARD;      PRT;     5 AA.
AC   P38005;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Unknown protein from 2D-page from elementary body (Fragment).
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=L2/434/BU;
RA   Bini L., Santucci A., Magi B., Marzocchi B., Sanchez-Campillo M.,
RA   Comanducci M., Christianen G., Birkelund S., Vtretou E., Ratti G.,
RA   Pallini V.;
RL   Submitted (SEP-1994) to the SWISS-PROT data bank.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 4.5, ITS MW IS: 28 kDa.
DR   Siena-2DPAGE; P38005; -.
FT   NON_TER       5      5
SQ   SEQUENCE   5 AA;  474 MW;  75BAA865AA800000 CRC64;


  Query Match             18.5%;  Score 12;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 GD 6
          ||
Db      3 GD 4


RESULT  12
ACI_THUAL
ID   ACI_THUAL      STANDARD;      PRT;     8 AA.
AC   P18691;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-NOV-1990 (Rel. 16, Last annotation update)
```

```
DE   Angiotensin-converting enzyme inhibitor.
OS   Thunnus albacares (Yellowfin tuna) (Neothunnus macropterus).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC   Scombridae; Thunnus.
OX   NCBI_TaxID=8236;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Muscle;
RX   MEDLINE=88326322; PubMed=3415688;
RA   Kohama Y., Matsumoto S., Oka H., Teramoto T., Okabe M., Mimura T.;
RT   "Isolation of angiotensin-converting enzyme inhibitor from tuna
RT   muscle.";
RL   Biochem. Biophys. Res. Commun. 155:332-337(1988).
DR   PIR; A31570; A31570.
SQ   SEQUENCE   8 AA;  953 MW;  6AA863733051F1B7 CRC64;

  Query Match             18.5%;  Score 12;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 GD 6
          ||
Db      7 GD 8


RESULT  13
LMT2_LOCMI
ID   LMT2_LOCMI     STANDARD;      PRT;     8 AA.
AC   P22396;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-AUG-1991 (Rel. 19, Last annotation update)
DE   Locustamyotropin 2 (LOM-MT-2).
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC   Acridomorpha; Acridoidea; Acrididae; Locusta.
OX   NCBI_TaxID=7004;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Corpora cardiaca;
RA   Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., de Loof A.;
RT   "Isolation, identification and synthesis of locustamyotropin II, an
RT   additional neuropeptide of Locusta migratoria. Member of the
RT   cephalomyotropic peptide family.";
RL   Insect Biochem. 20:479-484(1990).
CC   -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC       (MYOTROPIC ACTIVITY).
CC   -!- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
DR   InterPro; IPR001484; Pyrokinin.
DR   PROSITE; PS00539; PYROKININ; 1.
KW   Neuropeptide; Amidation; Pyrokinin.
FT   MOD_RES       8       8       AMIDATION.
SQ   SEQUENCE   8 AA;  934 MW;  26341771A9CAA87B CRC64;

  Query Match             18.5%;  Score 12;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 GD 6
          ||
Db      2 GD 3


RESULT  14
ORMY_ORCLI
ID   ORMY_ORCLI     STANDARD;      PRT;     8 AA.
AC   P82455;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Orcomyotropin (OMT).
OS   Orconectes limosus (Spinycheek crayfish).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
OC   Astacoidea; Cambaridae; Orconectes.
OX   NCBI_TaxID=28379;
RN   [1]
RP   SEQUENCE, MASS SPECTROMETRY, AND AMIDATION.
RC   TISSUE=Hindgut;
RX   MEDLINE=20411310; PubMed=10952880;
RA   Dircksen H., Burdzik S., Sauter A., Keller R.;
RT   "Two orcokinins and the novel octapeptide orcomyotropin in the hindgut
RT   of the crayfish Orconectes limosus: identified myostimulatory
RT   neuropeptides originating together in neurones of the terminal
RT   abdominal ganglion.";
RL   J. Exp. Biol. 203:2807-2818(2000).
CC   -!- FUNCTION: MYOTROPIC PEPTIDE, ENHANCES BOTH THE FREQUENCY AND
CC       AMPLITUDE OF SPONTANEOUS HINDGUT CONTRACTIONS. IT IS SYNTHESIZED
CC       BY ABDOMINAL GANGLIONIC NEURONS.
CC   -!- MASS SPECTROMETRY: MW=904.8; METHOD=FAB.
KW   Amidation; Neuropeptide.
FT   MOD_RES       8       8       AMIDATION.
SQ   SEQUENCE   8 AA;  905 MW;  87C861B1A9CDDAA9 CRC64;

  Query Match             18.5%;  Score 12;  DB 1;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 1e+05;
  Matches    2;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 DCF 8
          | |
Db      2 DAF 4


RESULT  15
ISOT_CYPCA
ID   ISOT_CYPCA     STANDARD;      PRT;     9 AA.
AC   P42993;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Isotocin.
OS   Cyprinus carpio (Common carp).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC   Cypriniformes; Cyprinidae; Cyprinus.
OX   NCBI_TaxID=7962;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Pituitary;
RA   Acher R., Chauvet J., Chauvet M.-T., Crepy D.;
RT   "Characterization of neurohypophyseal hormones from a fresh water bony
RT   fish, the carp (Cyprinus carpio). Comparison with hormones from sea
RT   water bony fishs.";
RL   Comp. Biochem. Physiol. 14:245-254(1965).
CC   -!- FUNCTION: ANTIDIURETIC HORMONE.
CC   -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR   PIR; A61364; A61364.
DR   InterPro; IPR000981; Neurhypophys_horm.
DR   Pfam; PF00220; hormone4; 1.
DR   PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW   Hormone; Amidation.
FT   DISULFID      1       6
FT   MOD_RES       9       9       AMIDATION.
SQ   SEQUENCE   9 AA;  969 MW;  17FF476EB455B04B CRC64;

  Query Match             18.5%;  Score 12;  DB 1;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1e+05;
  Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

Qy 7 CF 8
|:
Db 1 CY 2

Search completed: May 29, 2002, 10:01:26
Job time: 225 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 29, 2002, 09:56:56 ; Search time 25.79 Seconds
(without alignments)
60.370 Million cell updates/sec

Title: US-09-734-628-1
Perfect score: 65
Sequence: 1 CDCRGDCFC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 648

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	24	36.9	8	6	Q9TRY3	Q9try3 sus sp. ins
2	22	33.8	9	15	O12096	O12096 caprine art
3	22	33.8	9	15	O12098	O12098 caprine art
4	22	33.8	9	15	O12100	O12100 caprine art
5	22	33.8	9	15	O12102	O12102 caprine art
6	22	33.8	9	15	O12104	O12104 caprine art
7	19.5	30.0	9	10	Q9FSZ2	Q9fsz2 cicer ariet
8	19.5	30.0	9	11	Q99JF4	Q99jf4 mus musculu
9	19	29.2	7	11	O55184	O55184 rattus norv
10	17	26.2	8	4	Q9BYY5	Q9byy5 homo sapien
11	17	26.2	8	6	Q9BFC3	Q9bfc3 didelphis m
12	17	26.2	8	6	Q9BFC2	Q9bfc2 macropus eu
13	17	26.2	8	6	Q9BFC1	Q9bfc1 choloepus h
14	17	26.2	8	6	Q9BFC0	Q9bfc0 choloepus d
15	17	26.2	8	6	Q9BFB9	Q9bfb9 euphractus
16	17	26.2	8	6	Q9BFB8	Q9bfb8 chaetophrac
17	17	26.2	8	6	Q9BFB7	Q9bfb7 tamandua te
18	17	26.2	8	6	Q9BFB6	Q9bfb6 myrmecophag
19	17	26.2	8	6	Q9BFB5	Q9bfb5 erinaceus c
20	17	26.2	8	6	Q9BFB4	Q9bfb4 talpa altai
21	17	26.2	8	6	Q9BFB3	Q9bfb3 condylura c
22	17	26.2	8	6	Q9BFB2	Q9bfb2 sorex arane
23	17	26.2	8	6	Q9BFB1	Q9bfb1 echinops te
24	17	26.2	8	6	Q9BFB0	Q9bfb0 trichechus
25	17	26.2	8	6	Q9BFA9	Q9bfa9 procavia ca
26	17	26.2	8	6	Q9BFA8	Q9bfa8 loxodonta a
27	17	26.2	8	6	Q9BFA6	Q9bfa6 orycteropus
28	17	26.2	8	6	Q9BFA5	Q9bfa5 cynocephalu
29	17	26.2	8	6	Q9BFA4	Q9bfa4 tupaia mino
30	17	26.2	8	6	Q9BFA3	Q9bfa3 lemur catta
31	17	26.2	8	6	Q9BFA2	Q9bfa2 tarsius ban
32	17	26.2	8	6	Q9BFA1	Q9bfa1 ateles fusc
33	17	26.2	8	6	Q9BFA0	Q9bfa0 macaca mula
34	17	26.2	8	6	Q9BF99	Q9bf99 hylobates c
35	17	26.2	8	6	Q9BF98	Q9bf98 callimico g
36	17	26.2	8	6	Q9BF97	Q9bf97 artibeus ja
37	17	26.2	8	6	Q9BF96	Q9bf96 pteropus gi
38	17	26.2	8	6	Q9BF95	Q9bf95 rousettus l
39	17	26.2	8	6	Q9BF94	Q9bf94 nycteris th
40	17	26.2	8	6	Q9BF93	Q9bf93 megaptera n
41	17	26.2	8	6	Q9BF92	Q9bf92 tursiops tr
42	17	26.2	8	6	Q9BF91	Q9bf91 hippopotamu
43	17	26.2	8	6	Q9BF90	Q9bf90 tragelaphus
44	17	26.2	8	6	Q9BF89	Q9bf89 okapia john
45	17	26.2	8	6	Q9BF88	Q9bf88 equus cabal

ALIGNMENTS

RESULT 1
Q9TRY3
ID Q9TRY3 PRELIMINARY; PRT; 8 AA.
AC Q9TRY3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE INSULIN-LIKE GROWTH FACTOR-BINDING PROTEIN-6, IGFBP-6.
OS Sus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9826;
RN [1]
RP SEQUENCE.
RX MEDLINE=92049376; PubMed=1719383;
RA Shimasaki S., Gao L., Shimonaka M., Ling N.;
RT "Isolation and molecular cloning of insulin-like growth factor-binding
RT protein-6.";
RL Mol. Endocrinol. 5:938-948(1991).
SQ SEQUENCE 8 AA; 850 MW; 9FB2CEA37EA7687D CRC64;

Query Match 36.9%; Score 24; DB 6; Length 8;
Best Local Similarity 60.0%; Pred. No. 5.6e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy      5 GDCFC 9
          | |:|
Db      2 GPCWC 6
```

RESULT 2
O12096
ID O12096 PRELIMINARY; PRT; 9 AA.
AC O12096;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE TAT PROTEIN (FRAGMENT).

```
GN   TAT.
OS   Caprine arthritis encephalitis virus (CAEV).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11660;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;
RT   "dUTPase minus CAEV is attenuated for pathogenesis and accumulates G
RT   to A substitutions.";
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U81439; AAB60832.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   9 AA;  922 MW;  21E8644EB7340EB8 CRC64;


  Query Match             33.8%;  Score 22;  DB 15;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CDCR 4
          | ||
Db      1 CGCR 4


RESULT   3
O12098,
ID   O12098      PRELIMINARY;      PRT;     9 AA.
AC   O12098;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TAT PROTEIN (FRAGMENT).
GN   TAT.
OS   Caprine arthritis encephalitis virus (CAEV).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11660;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;
RT   "dUTPase minus CAEV is attenuated for pathogenesis and accumulates G
RT   to A substitutions.";
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U81440; AAB60835.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   9 AA;  922 MW;  21E8644EB7340EB8 CRC64;


  Query Match             33.8%;  Score 22;  DB 15;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CDCR 4
          | ||
Db      1 CGCR 4


RESULT   4
O12100
ID   O12100      PRELIMINARY;      PRT;     9 AA.
AC   O12100;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TAT PROTEIN (FRAGMENT).
GN   TAT.
OS   Caprine arthritis encephalitis virus (CAEV).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11660;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;
RT   "dUTPase minus CAEV is attenuated for pathogenesis and accumulates G
RT   to A substitutions.";
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U81441; AAB60836.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   9 AA;  922 MW;  21E8644EB7340EB8 CRC64;


  Query Match             33.8%;  Score 22;  DB 15;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CDCR 4
          | ||
Db      1 CGCR 4


RESULT   5
O12102
ID   O12102      PRELIMINARY;      PRT;     9 AA.
AC   O12102;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TAT PROTEIN (FRAGMENT).
GN   TAT.
OS   Caprine arthritis encephalitis virus (CAEV).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11660;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;
RT   "dUTPase minus CAEV is attenuated for pathogenesis and accumulates G
RT   to A substitutions.";
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U81442; AAB60838.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   9 AA;  922 MW;  21E8644EB7340EB8 CRC64;


  Query Match             33.8%;  Score 22;  DB 15;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CDCR 4
          | ||
Db      1 CGCR 4


RESULT   6
O12104
ID   O12104      PRELIMINARY;      PRT;     9 AA.
AC   O12104;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TAT PROTEIN (FRAGMENT).
GN   TAT.
OS   Caprine arthritis encephalitis virus (CAEV).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11660;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;
RT   "dUTPase minus CAEV is attenuated for pathogenesis and accumulates G
RT   to A substitutions.";
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U81443; AAB60840.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   9 AA;  922 MW;  21E8644EB7340EB8 CRC64;


  Query Match             33.8%;  Score 22;  DB 15;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy      1 CDCR 4
          | ||
Db      1 CGCR 4


RESULT   7
Q9FSZ2
ID   Q9FSZ2      PRELIMINARY;      PRT;     9 AA.
AC   Q9FSZ2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   HYPOTHETICAL 1.0 KDA PROTEIN (FRAGMENT).
OS   Cicer arietinum (Chickpea) (Garbanzo).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae; Cicer.
OX   NCBI_TaxID=3827;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. CASTELLANA; TISSUE=ETIOLATED EPICOTYLS;
RA   Dopico B., Jimenez T., Labrador E.;
RT   "cDNA clones expressed in etiolated Cicer arietinum epicotyls.";
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ299069; CAC10216.1; -.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  990 MW;  9441BDDAA7272EBE CRC64;


  Query Match             30.0%;  Score 19.5;  DB 10;  Length 9;
  Best Local Similarity   55.6%;  Pred. No. 5.6e+05;
  Matches    5;  Conservative    0;  Mismatches    3;  Indels    1;  Gaps    1;

Qy      1 CDCRGD-CF 8
          | |  | ||
Db      1 CCCLLDACF 9


RESULT   8
Q99JF4
ID   Q99JF4      PRELIMINARY;      PRT;     9 AA.
AC   Q99JF4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   OCT-1L (FRAGMENT).
GN   OCT-1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Pankratova E.V., Deyev I.E., Zhenilo S.V., Polanovsky O.L.;
RT   "Tissue-specific Oct-1 isoforms from murine lymphocytes.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ310124; CAC34946.1; -.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  998 MW;  540BCEBAB5BEBAA7 CRC64;


  Query Match             30.0%;  Score 19.5;  DB 11;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 5.6e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy      2 DCRGDC 7
          ||  ||
Db      3 DC-SDC 7


RESULT   9
O55184
ID   O55184      PRELIMINARY;      PRT;     7 AA.
AC   O55184;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ORPHAN RECEPTOR TR4-NS (FRAGMENT).
GN   TR4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY;
RX   MEDLINE=96198747; PubMed=8612486;
RA   Yoshikawa T., Makino S., Gao X.M., Xing G.Q., Chuang D.M.,
RA   Detera-Wadleigh S.D.;
RT   "Splice variants of rat TR4 orphan receptor: differential expression
RT   of novel sequences in the 5'-untranslated region and C-terminal
RT   domain.";
RL   Endocrinology 137:1562-1571(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY;
RX   MEDLINE=96299786; PubMed=8661150;
RA   Yoshikawa T., DuPont B.R., Leach R.J., Detera-Wadleigh S.D.;
RT   "New variants of the human and rat nuclear hormone receptor, TR4:
RT   expression and chromosomal localization of the human gene.";
RL   Genomics 35:361-366(1996).
DR   EMBL; U59454; AAB91433.1; -.
KW   Receptor.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  663 MW;  6DDAA8787EB05350 CRC64;


  Query Match             29.2%;  Score 19;  DB 11;  Length 7;
  Best Local Similarity   75.0%;  Pred. No. 5.6e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 CRGD 6
          | ||
Db      3 CGGD 6


RESULT  10
Q9BYY5
ID   Q9BYY5      PRELIMINARY;      PRT;     8 AA.
AC   Q9BYY5;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CAMP RESPONSIVE ELEMENT MODERATOR (FRAGMENT).
GN   CREM.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21082082; PubMed=11214319;
RA   Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA   O'Brien S.J.;
RT   "Molecular phylogenetics and the origins of placental mammals.";
RL   Nature 409:614-618(2001).
DR   EMBL; AY011664; AAG47575.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   8 AA;  1006 MW;  DF02C331EEAB572A CRC64;


  Query Match             26.2%;  Score 17;  DB 4;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy      6 DCFC 9
          | :|
Db      1 DLYC 4


RESULT  11
Q9BFC3
ID   Q9BFC3      PRELIMINARY;      PRT;     8 AA.
AC   Q9BFC3;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CAMP RESPONSIVE ELEMENT MODERATOR (FRAGMENT).
GN   CREM.
OS   Didelphis marsupialis virginiana (North American opossum).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.
OX   NCBI_TaxID=9267;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21082082; PubMed=11214319;
RA   Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA   O'Brien S.J.;
RT   "Molecular phylogenetics and the origins of placental mammals.";
RL   Nature 409:614-618(2001).
DR   EMBL; AY011620; AAG47535.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   8 AA;  978 MW;  DF1DD331EEAB572A CRC64;

  Query Match             26.2%;  Score 17;  DB 6;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 DCFC 9
          | :|
Db      1 DLYC 4


RESULT  12
Q9BFC2
ID   Q9BFC2      PRELIMINARY;      PRT;     8 AA.
AC   Q9BFC2;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CAMP RESPONSIVE ELEMENT MODERATOR (FRAGMENT).
GN   CREM.
OS   Macropus eugenii (Tammar wallaby).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX   NCBI_TaxID=9315;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21082082; PubMed=11214319;
RA   Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA   O'Brien S.J.;
RT   "Molecular phylogenetics and the origins of placental mammals.";
RL   Nature 409:614-618(2001).
DR   EMBL; AY011621; AAG47536.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   8 AA;  978 MW;  DF1DD331EEAB572A CRC64;

  Query Match             26.2%;  Score 17;  DB 6;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 DCFC 9
          | :|
Db      1 DLYC 4


RESULT  13
Q9BFC1
ID   Q9BFC1      PRELIMINARY;      PRT;     8 AA.
AC   Q9BFC1;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CAMP RESPONSIVE ELEMENT MODERATOR (FRAGMENT).
GN   CREM.
OS   Choloepus hoffmanni (Hoffmann's two-fingered sloth).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Edentata; Choloepidae; Choloepus.
OX   NCBI_TaxID=9358;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21082082; PubMed=11214319;
RA   Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA   O'Brien S.J.;
RT   "Molecular phylogenetics and the origins of placental mammals.";
RL   Nature 409:614-618(2001).
DR   EMBL; AY011622; AAG47537.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   8 AA;  978 MW;  DF1DD331EEAB572A CRC64;

  Query Match             26.2%;  Score 17;  DB 6;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 DCFC 9
          | :|
Db      1 DLYC 4


RESULT  14
Q9BFC0
ID   Q9BFC0      PRELIMINARY;      PRT;     8 AA.
AC   Q9BFC0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CAMP RESPONSIVE ELEMENT MODERATOR (FRAGMENT).
GN   CREM.
OS   Choloepus didactylus (southern two-toed sloth).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Edentata; Choloepidae; Choloepus.
OX   NCBI_TaxID=27675;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21082082; PubMed=11214319;
RA   Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA   O'Brien S.J.;
RT   "Molecular phylogenetics and the origins of placental mammals.";
RL   Nature 409:614-618(2001).
DR   EMBL; AY011623; AAG47538.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   8 AA;  978 MW;  DF1DD331EEAB572A CRC64;

  Query Match             26.2%;  Score 17;  DB 6;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 DCFC 9
          | :|
Db      1 DLYC 4


RESULT  15
Q9BFB9
ID   Q9BFB9      PRELIMINARY;      PRT;     8 AA.
AC   Q9BFB9;
```

```
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CAMP RESPONSIVE ELEMENT MODERATOR (FRAGMENT).
GN   CREM.
OS   Euphractus sexcinctus (Six-banded armadillo).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Edentata; Dasypodidae; Euphractus.
OX   NCBI_TaxID=143300;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21082082; PubMed=11214319;
RA   Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA   O'Brien S.J.;
RT   "Molecular phylogenetics and the origins of placental mammals.";
RL   Nature 409:614-618(2001).
DR   EMBL; AY011624; AAG47539.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   8 AA;  978 MW;  DF1DD331EEAB572A CRC64;


  Query Match             26.2%;  Score 17;  DB 6;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 5.6e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 DCFC 9
           | :|
Db       1 DLYC 4


Search completed: May 29, 2002, 10:01:03
Job time: 247 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 29, 2002, 10:41:42 ; Search time 28.88 Seconds
(without alignments)
34.614 Million cell updates/sec

Title: US-09-734-628-1
Perfect score: 65
Sequence: 1 CDCRGDCFC 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 102553

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	65	100.0	9	16	AAR76200	Alphav/beta3 and a
2	65	100.0	9	19	AAW60289	Tumour homing pept
3	65	100.0	9	19	AAW56034	Chimeric adenoviru
4	65	100.0	9	20	AAY43233	RGD-containing pep
5	65	100.0	9	20	AAY48821	Membrane dipeptida
6	65	100.0	9	20	AAY42255	Synthetic RGD-4C p
7	65	100.0	9	20	AAW93626	NGR receptor bindi
8	65	100.0	9	21	AAB21701	Human breast tumou
9	65	100.0	9	21	AAB17346	Integrin-binding p
10	65	100.0	9	21	AAB17928	TPO-mimetic peptid
11	65	100.0	9	21	AAB17964	Integrin-binding p
12	65	100.0	9	21	AAY90211	Alphav integrin ta
13	65	100.0	9	21	AAY44970	RGD-4C targeting s
14	65	100.0	9	21	AAY54271	Alpha Vbeta-3 bind
15	65	100.0	9	22	AAE11044	RGD-containing pep
16	65	100.0	9	22	AAE06279	Tumour homing pept
17	65	100.0	9	22	AAB97086	Integrin-binding p
18	65	100.0	9	22	AAB20271	Peptide that speci
19	65	100.0	9	22	AAB50242	Enhanced infectivi
20	59	90.8	9	16	AAR79073	Alphav/beta3 and a
21	59	90.8	9	21	AAB17347	Integrin-binding p
22	51	78.5	9	16	AAR76199	Alphav/beta3 and a
23	51	78.5	9	19	AAW56035	Chimeric adenoviru
24	51	78.5	9	21	AAB17345	Integrin-binding p
25	49	75.4	9	16	AAR79074	Alphav/beta3 and a
26	49	75.4	9	21	AAB17348	Integrin-binding p
27	47	72.3	7	21	AAY90212	Alphav integrin ta
28	47	72.3	7	21	AAY90219	UPAR targeting seq
29	44	67.7	7	20	AAY43232	RGD-containing pep
30	40	61.5	8	16	AAR96333	RGD cyclic peptide
31	40	61.5	8	19	AAW66836	Peptide useful for
32	40	61.5	8	20	AAW97017	Peptide used to in
33	35	53.8	5	12	AAR10414	Fibrinogen recepto
34	35	53.8	5	12	AAR10415	Fibrinogen recepto
35	35	53.8	5	12	AAR10418	Fibrinogen recepto
36	35	53.8	5	12	AAR11587	Fibrinogen recepto
37	35	53.8	5	13	AAR27031	Peptide lipid cont
38	35	53.8	5	13	AAR29052	Peptide lipid cont
39	35	53.8	5	14	AAR69325	Gp IIb/IIIa recept
40	35	53.8	5	16	AAR79093	Alpha5/beta1 integ
41	35	53.8	5	17	AAW03492	Alpha(5)-Beta(1) i
42	35	53.8	5	19	AAW64952	Targetting peptide
43	35	53.8	5	19	AAW48499	Integrin receptor
44	35	53.8	5	19	AAW50594	GPIIb/IIIa recepto
45	35	53.8	5	20	AAY21570	Integrin-binding p

ALIGNMENTS

RESULT 1
AAR76200
ID AAR76200 standard; peptide; 9 AA.
XX
AC AAR76200;
XX
DT 24-JAN-1996 (first entry)
XX
DE Alphav/beta3 and alphav/beta5 integrin binding peptide #4.
XX
KW High affinity; integrin binding peptide; alpha5/beta1; alphav/beta5;
KW alphav/beta3; RGD; stable configuration; wound healing;
KW osteoclast attachemnt; bone; angiogenesis; metastasis; tumour;
KW smooth muscle cell migration.
XX
OS Synthetic.
XX
PN WO9514714-A1.
XX
PD 01-JUN-1995.
XX
PF 22-NOV-1994; 94WO-US13542.
XX
PR 04-AUG-1994; 94US-0286861.
PR 24-NOV-1993; 93US-0158001.
XX
PA (LJOL-) LA JOLLA CANCER RES FOUND.
XX
PI Koivunen E, Ruoslahti E;
XX
DR WPI; 1995-206899/27.
XX
PT High affinity integrin binding peptides - can be used to attach
PT cells to a substrate, inhibit the attachment of osteoclasts to bone,

PT promote wound healing, inhibit angiogenesis, metastasis of tumours
PT and migration of smooth muscle cells
XX
PS Claim 21; Page 62; 86pp; English.
XX
CC The sequences given in AAR76185-200 and AAR79073-94 are high affinity
CC integrin binding peptides which bind to various integrins. Peptides
CC which bind to alpha5/beta1 integrins contain the motifs given in
CC AAR76185-86 and peptides which bind to alphav/beta5 and alphav/beta3
CC integrins contain the motif given in AAR76187. Alphav/beta5 integrins
CC are also bound by RGD containing peptides. These peptides assume a
CC conformationally stabilised configuration which is due to the
CC formation of a disulphide bond, a peptide bond or a lactam bond.
CC These peptides may be used for isolating the complementary integrin
CC from a sample mixture by contacting them under ionic conditions to
CC allow binding of the integrin to the peptide and then separating the
CC integrin from the peptide. They can be used for attaching cells to
CC a substrate, by binding them to the substrate with the cell. The
CC peptides promote wound healing when applied locally and inhibit the
CC attachment of osteoclasts to bone. They inhibit angiogenesis,
CC metastasis of tumours and migration of smooth muscle cells.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 2
AAW60289
ID AAW60289 standard; peptide; 9 AA.
XX
AC AAW60289;
XX
DT 24-AUG-1998 (first entry)
XX
DE Tumour homing peptide of the invention.
XX
KW Tumour homing peptide; in vivo panning;
KW alpha-V-containing integrin binding motif; tumour.
XX
OS Unidentified.
XX
PN WO9810795-A2.
XX
PD 19-MAR-1998.
XX
PF 10-SEP-1997; 97WO-US16086.
XX
PR 10-SEP-1996; 96US-0710067.
XX
PA (BURN-) BURNHAM INST.
XX
PI Pasqualini R, Ruoslahti E;
XX
DR WPI; 1998-207151/18.
XX
PT Tumour homing molecules and their conjugates - useful for, e g.
PT directing linked moiety to tumour containing angiogenic vasculature
XX
PS Claim 6; Page 91; 105pp; English.
XX
CC The present peptide represents a tumour homing peptide, and is produced
CC by in vivo panning. The peptide has an alpha-V-containing integrin
CC binding motif, Arg-Gly-Asp (RGD). The in vivo panning comprises
CC administering a library of diverse peptides to a subject having a
CC tumour, collecting a sample of the tumour, identifying a peptide that
CC homes to the tumour, collecting a sample of normal tissue corresponding
CC to the tumour, and determining that the peptide that homes to the
CC tumour is not present in the normal tissue. The tumour homing peptide can
CC be linked to a moiety (e.g. doxorubicin), and used to direct the
CC moiety to a tumour.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 19; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 3
AAW56034
ID AAW56034 standard; peptide; 9 AA.
XX
AC AAW56034;
XX
DT 29-JUL-1998 (first entry)
XX
DE Chimeric adenovirus fiber protein non-native amino acid sequence 3.
XX
KW Chimeric; adenovirus; fiber protein; binding; targeting; coat protein;
KW constrained peptide motif; gene therapy; cancer; heart disease;
KW autoimmune disorder.
XX
OS Synthetic.
OS Mastadenovirus.
XX
PN WO9807865-A1.
XX
PD 26-FEB-1998.
XX
PF 21-AUG-1997; 97WO-US14719.
XX
PR 21-AUG-1996; 96US-0701124.
XX
PA (GENV-) GENVEC INC.
XX
PI Kovesdi I, Roelvink PW, Wickham TJ;
XX
DR WPI; 1998-169169/15.
XX
PT Chimeric adenovirus fibre proteins - containing non-native amino
PT acid sequence to provide for binding and entry into cells,
PT especially for gene therapy
XX
PS Claim 7; Page 68; 124pp; English.
XX
CC The present sequence represents a specifically claimed non-native amino
CC acid sequence from a chimeric adenovirus fibre protein (AFP) of the
CC present invention. The non-native amino acid sequence allows the
CC chimeric fibre (or a vector comprising the chimeric fibre) to more
CC efficiently bind to and enter cells. The products can be used for gene
CC therapy, for treating cancer, e.g. melanoma, glioma and lung cancers as
CC well as genetic disorders, e.g. cystic fibrosis, haemophilia and
CC muscular dystrophy as well as pathogenic infections, e.g. HIV,
CC tuberculosis and hepatitis and also for heart disease, to e.g. prevent
CC restenosis following angioplasty or to promote angiogenesis to reperfuse
CC necrotic tissue, and in autoimmune disorders, e.g. Crohn's disease,
CC colitis, rheumatoid arthritis, and Alzheimer's disease.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 19; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 CDCRGDCFC 9
          |||||||||
Db      1 cdcrgdcfc 9
```

RESULT 4
AAY43233
ID AAY43233 standard; peptide; 9 AA.
XX
AC AAY43233;
XX
DT 13-JAN-2000 (first entry)
XX
DE RGD-containing peptide #12.
XX
KW Nucleic acid delivery vehicle; bifunctional complex; transgene; CFTR;
KW cell surface targeting; cell surface molecule binding region; integrin;
KW cystic fibrosis transmembrane regulator; alphal-antitrypsin;
KW suicide gene; beta-glucocerebrosidase; cell transfection; cell infection;
KW RGD peptide.
XX
OS Synthetic.
XX
PN WO9940214-A2.
XX
PD 12-AUG-1999.
XX
PF 08-FEB-1999; 99WO-US02680.
XX
PR 09-FEB-1998; 98US-0020483.
PR 06-NOV-1998; 98US-0107471.
XX
PA (GENZ) GENZYME CORP.
XX
PI O'riordan C, Romanczuk H, Wadsworth SC;
XX
DR WPI; 1999-610583/52.
XX
PT Nucleic acid delivery vehicles useful for transfecting and infecting a
PT target cell -
XX
PS Claim 22; Page 39; 118pp; English.
XX
CC This sequence represents a RGD-containing peptide that can be used in a
CC bifunctional complex used in the nucleic acid delivery vehicle (I) of the
CC invention. (I) is for transfecting and/or infecting a target cell, and
CC comprises a transgene and a bifunctional complex (B) that targets the
CC nucleic acid delivery vehicle to the cell surface. (B) comprises a
CC delivery vehicle binding portion, a cell surface molecule binding portion
CC (such as this sequence) and a linker connecting them. The delivery
CC vehicle can be specifically targeted to the cell via the binding to cell
CC surface molecules. (I) can be used to target cells, which express
CC integrins such as, HT-29 colon carcinoma cells, lymphocytes and
CC monocytes, blood platelets, SMC-90 human lung fibroblast, MG(63)
CC osteosarcoma cell line, vascular endothelial cells and melanoma cells.
CC (I) is useful for delivery of nucleic acids encoding CFTR (cystic
CC fibrosis transmembrane regulator), alphal-antitrypsin,
CC beta-glucocerebrosidase and suicide genes. The construct increases the
CC efficiency of cellular uptake of (I). The constructs also enable the
CC transfection/infection of cells that are normally refractory to
CC transfection/infection by targeting cell receptors that are present on
CC such cells.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 CDCRGDCFC 9
          |||||||||
Db      1 cdcrgdcfc 9
```

RESULT 5
AAY48821
ID AAY48821 standard; Peptide; 9 AA.
XX
AC AAY48821;
XX
DT 10-DEC-1999 (first entry)
XX
DE Membrane dipeptidase-binding retina homing peptide #7.
XX
KW Homing peptide; organ; tissue; lung; pancreas; skin; retina; MDP;
KW prostate; ovary; lymph node; adrenal gland; liver; gut; tumour;
KW membrane dipeptidase.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9946284-A2.
XX
PD 16-SEP-1999.
XX
PF 10-MAR-1999; 99WO-US05284.
XX
PR 13-MAR-1998; 98US-0042107.
PR 26-FEB-1999; 99US-0042107.
XX
PA (BURN-) BURNHAM INST.
XX
PI Rajotte D, Pasqualini R, Ruoslahti EI;
XX
DR WPI; 1999-571717/48.
XX
PT New peptides which selectively home to organs or tissues, used for,
PT e.g. identifying target ligands and for therapy of pathological
PT conditions -
XX
PS Example 6; Page 149; 193pp; English.
XX
CC The present invention describes peptides that selectively home to a
CC tissue or organ. The peptides can be used for identifying an organ
CC or tissue, for identifying a target molecule expressed by an organ or
CC tissue or for treating an organ or tissue pathology, where the organ or
CC tissue is selected from prostate, lung, skin, retina, pancreas, gut,
CC ovary, adrenal gland, liver, and lymph node. The peptide bind to the
CC membrane dipeptidase (MDP). AAY48618 to AAY49066 represent sequences
CC which are used in the exemplification of the present invention.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 CDCRGDCFC 9
          |||||||||
Db      1 cdcrgdcfc 9
```

RESULT 6
AAY42255
ID AAY42255 standard; peptide; 9 AA.
XX
AC AAY42255;
XX
DT 01-DEC-1999 (first entry)
XX
DE Synthetic RGD-4C peptide.
XX

KW Adenovirus; gene therapy; coxsackievirus adenovirus receptor;
KW CAR; cancer; cystic fibrosis; muscular dystrophy.
XX
OS Synthetic.
XX
PN WO9939734-A1.
XX
PD 12-AUG-1999.
XX
PF 05-FEB-1999; 99WO-US02549.
XX
PR 06-FEB-1998; 98US-0073947.
PR 10-SEP-1998; 98US-0099801.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Curiel DT, Krasnykh VN, Dmitriev I;
XX
DR WPI; 1999-539951/45.
XX
Recombinant adenovirus vectors with modified fiber knob loops, useful
in gene therapy -
XX
PS Example 21; Page 49; 126pp; English.
XX
CC This sequence represents a synthetic RGD-4C peptide. DNA encoding
CC this sequence was cloned into the sequence encoding the HI loop of the
CC adenovirus fibre protein knob domain. This was then used in the
CC construction of plasmids encoding a modified fibre protein. Recombinant
CC adenovirus genomes were generated by homologous DNA recombination in E.
CC coli, before excision of the newly generated genome for virus rescue.
CC The knob domain of the adenovirus fibre protein mediates the initial
CC binding and recognition of the coxsackievirus and adenovirus receptor
CC (CAR) on the cell surface. The HI loop protrudes from the knob domain
CC and connects beta-strands involved in the formation of the cell binding
CC site. Recombinant adenovirus vectors are used in a number of gene
CC therapy applications; however, the reliance on the CAR means that
CC in certain situations, recombinant virions are sequestered by high
CC CAR-expressing non-target cells while the true target cells, if low
CC in CAR, receive little of the therapeutic gene. Modification of the HI
CC loop by replacement of the hypervariable region of the loop with a
CC peptide such as the RGD peptide results in the
CC ability of the virus to utilise an alternative receptor during the cell
CC entry process. Modifying the adenovirus fibre knob protein in this way
CC increases the ability of an adenovirus to transduce a tumour cell in
CC vitro, in vivo and ex vivo. The vector Ad5FHIFLAG incorporating an RGD
CC peptide demonstrated two to three orders of magnitude
CC of increased gene transfer to ovarian cancer cells. The modified
CC adenovirus has an altered tropism, which allows the adenovirus to be
CC targeted to selected cell types. The recombinant adenovirus can be used
CC to provide gene therapy for individuals suffering from cancer, cystic
CC fibrosis and Duchenne's muscular dystrophy.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 7
AAW93626
ID AAW93626 standard; Protein; 9 AA.
XX
AC AAW93626;
XX
DT 28-JUN-1999 (first entry)
XX
DE NGR receptor binding tumour homing peptide 5.
XX
KW Tumour homing peptide; tumour; diagnosis; endothelial cell;
KW angiogenic vasculature; anti-tumour; anti-inflammatory; anti-angiogenic;
KW anti-arthritic; NGR receptor; inhibitor; angiogenesis; anticancer drug;
KW prognosis; inflammation; regeneration; wounded tissue; targetting;
KW macular degeneration; diabetic retinopathy; rheumatoid arthritis;
KW occlusive thrombus.
XX
OS Synthetic.
XX
PN WO9913329-A1.
XX
PD 18-MAR-1999.
XX
PF 08-SEP-1998; 98WO-US18895.
XX
PR 25-AUG-1998; 98US-0139802.
PR 10-SEP-1997; 97US-0926914.
XX
PA (BURN-) BURNHAM INST.
XX
PI Pasqualini R, Ruoslahti E;
XX
DR WPI; 1999-215158/18.
XX
PT Identifying molecules that home to angiogenic vasculature used as
PT targets for anticancer agents
XX
PS Claim 15; Page 7; 180pp; English.
XX
CC This invention describes novel peptides which home to angiogenic
CC vasculature, specifically of a tumour and which have anti-tumour,
CC anti-inflammatory, anti-angiogenic and anti-arthritic activity. Such
CC molecules are identified by treating a purified NGR receptor with a test
CC compound and identifying compounds that bind specifically to the NGR
CC receptor. The peptides of the invention are inhibitors of angiogenesis
CC and can be used to produce conjugates for delivering agents to
CC angiogenic vasculature, particularly anticancer drugs or an imaging
CC agent, for diagnosis or prognosis. These conjugates may be directed to
CC non-tumour angiogenic vasculature, e.g. that present in inflammatory,
CC regenerating or wounded tissue, e.g. for treatment of macular
CC degeneration, diabetic retinopathy or rheumatoid arthritis. The peptides
CC provide specific targeting to tumours, especially their supporting
CC vasculature, since the NGR receptor is exposed to the circulation only in
CC angiogenic vasculature. Precise targeting should reduce the systemic
CC toxicity of anticancer drugs in the conjugates. Complete killing of all
CC target cells may not be essential since partial denudation of endothelium
CC may result in an occlusive thrombus, and endothelial cells are unlikely
CC to become resistant to anticancer agents nor to lose the targeting
CC receptor. AAW93622-W93809 and AAW93843-44 are examples of tumour homing
CC peptides used in the invention.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 8
AAB21701
ID AAB21701 standard; Peptide; 9 AA.
XX
AC AAB21701;
XX
DT 22-MAR-2001 (first entry)
XX

DE Human breast tumour homing peptide #1.
XX
KW Cytostatic; homing pro-apoptotic conjugate; tumour; antimicrobial;
KW breast; prostate; melanoma; cancer; Kaposi's sarcoma; human.
XX
OS Homo sapiens.
XX
PN WO200042973-A2.
XX
PD 27-JUL-2000.
XX
PF 21-JAN-2000; 2000WO-US01602.
XX
PR 22-JAN-1999; 99US-0235902.
XX
PA (BURN-) BURNHAM INST.
XX
PI Ellerby HM, Bredesen DE, Pasqualini R, Ruoslahti EI;
XX
DR WPI; 2000-499174/44.
XX
PT Homing pro-apoptotic conjugate comprising a tumor homing molecule that
PT selectively homes to a mammalian cell type or tissue linked to an
PT antimicrobial peptide, useful for the treatment of prostate cancer -
XX
PS Claim 12; Page 105; 118pp; English.
XX
CC The present invention relates to homing pro-apoptotic conjugates,
CC comprising of a tumour homing molecule that selectively homes to a
CC mammalian cell type or tissue, linked to an antimicrobial peptide. The
CC homing pro-apoptotic conjugates are selectively internalised by the
CC mammalian cell type or tissue and exhibits high toxicity, especially to
CC angiogenic vasculature. The antimicrobial peptide has low mammalian cell
CC toxicity when not linked to the tumor homing molecule. The conjugates are
CC useful for the treatment of cancer e.g. Kaposi's sarcoma, breast and
CC prostate cancer or melanoma. The present sequence is a homing peptide
CC isolated in the present invention, which can be conjugated to an
CC antimicrobial peptide to make the homing pro-apoptotic conjugates of the
CC present invention.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 9
AAB17346
ID AAB17346 standard; Peptide; 9 AA.
XX
AC AAB17346;
XX
DT 31-OCT-2000 (first entry)
XX
DE Integrin-binding peptide sequence SEQ ID NO:450.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Claim 39; Page 354; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 10
AAB17928
ID AAB17928 standard; Peptide; 9 AA.
XX
AC AAB17928;
XX
DT 31-OCT-2000 (first entry)
XX
DE TPO-mimetic peptide sequence SEQ ID NO:1032.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX

PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Disclosure; Page 559; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 11
AAB17964
ID AAB17964 standard; Peptide; 9 AA.
XX
AC AAB17964;
XX
DT 31-OCT-2000 (first entry)
XX
DE Integrin-binding peptide sequence SEQ ID NO:1076.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Claim 39; Page 591; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 12
AAY90211
ID AAY90211 standard; peptide; 9 AA.
XX
AC AAY90211;
XX
DT 21-SEP-2000 (first entry)
XX
DE Alphav integrin targeting peptide #1.
XX
KW Ligand epitope; UPAR; urokinase-type plasminogen activator receptor;
KW adenovirus; hexon HVR5 loop; hexon HI loop; peripheral artery disease;
KW recombinant adenovirus vector; tumour; restenosis; gene therapy; asthma;
KW smooth muscle cell proliferation inhibitor; coronary artery disease;
KW obesity; neurodegenerative disease; infection; autoimmune disease; HIV;
KW thrombosis; diabetes; tropism-modified virus.
XX
OS Adenovirus sp.
XX
PN WO200012738-A1.
XX
PD 09-MAR-2000.
XX
PF 27-AUG-1999; 99WO-IB01524.

XX
PR 27-AUG-1998; 98US-0098028.
XX
PA (AVET) AVENTIS PHARMA SA.
XX
PI Vigne E, Dedieu J, Latta M, Yeh P, Perricaudet M;
XX
DR WPI; 2000-256653/22.
XX
PT Urokinase-type plasminogen activator receptor (UPAR)-targeted
PT adenovirus vectors having modified hexon HVR5 and HI loops and modified
PT fiber proteins useful for targeted gene therapy to treat cancer or
PT restenosis -
XX
PS Example 5; Page 53; 128pp; English.
XX
CC This sequence represents a alphav integrin targeting peptide.
CC The invention relates to an adenovirus from which at
CC least a part of the hexon HVR5 or HI loop is replaced with a binding
CC peptide, or targeting sequence, flanked by connecting amino acid spacers,
CC to functionally display its binding specificity at the capsid surface.
CC The invention also relates to a recombinant adenovirus vector where a
CC binding peptide, or targeting sequence, is connected to the C-terminus of
CC the fiber by a connecting spacer, or linker, so as to functionally
CC display its binding specificity at the capsid surface. The adenovirus or
CC recombinant adenovirus vector can be used to preferentially express a
CC gene in a target cell, especially a cell that expresses a UPAR. The
CC targeted adenovirus vector preferably comprises a heterologous gene
CC encoding a gene for treatment of a tumour or restenosis. The targeted
CC adenovirus vector is useful for gene therapy treatment of a disease, and
CC for manufacturing a medicine used in gene therapy treatment of a disease.
CC The viruses can also be used to inhibit smooth muscle cell proliferation,
CC to treat peripheral artery diseases, coronary artery diseases, obesity,
CC neurodegenerative diseases, infections, autoimmune diseases, asthma, HIV,
CC thrombosis, and diabetes. The viruses are particularly targeted against a
CC urokinase-type plasminogen activator receptor (UPAR). The adenoviruses
CC are tropism-modified without adversely impacting productivity of the
CC vectors.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 13
AAY44970
ID AAY44970 standard; Protein; 9 AA.
XX
AC AAY44970;
XX
DT 23-MAY-2000 (first entry)
XX
DE RGD-4C targeting sequence for KDEL receptor inhibitor protein.
XX
KW KDEL receptor inhibitor; heat shock protein; immune response;
KW oligomerisation domain; neoplasia; sarcoma; lymphoma; leukaemia;
KW melanoma; carcinoma; glioblastoma; astromytoma; oncogene;
KW infectious disease; allergy; autoimmune disease.
XX
OS Unidentified.
XX
PN WO200006729-A1.
XX
PD 10-FEB-2000.
XX
PF 28-JUL-1999; 99WO-US17147.
XX
PR 29-JUL-1998; 98US-0124671.
XX
PA (SLOK) SLOAN KETTERING INST CANCER RES.
XX
PI Rothman JE, Mayhew M, Hoe MH;
XX
DR WPI; 2000-195296/17.
XX
PT Inhibitors of the KDEL receptor which comprises an oligomerization
PT domain useful for promoting secretion of proteins which are normally
PT retained within the cell -
XX
PS Disclosure; Page 17; 87pp; English.
XX
CC The patent discloses the use of KDEL receptor inhibitor to promote
CC secretion of proteins that are normally retained within the cell such as
CC heat shock proteins by inhibiting KDEL receptor-mediated return of
CC protein complexes to endoplasmic reticulum. This makes the secreted heat
CC shock proteins more accessible to the immune system and improves immune
CC response to a target antigen. The inhibitor protein comprises several
CC subunits where each subunit comprises an oligomerisation domain and has
CC at its carboxy terminus a region which binds to a KDEL receptor. The
CC target antigen may be associated with diseases including neoplasia such
CC as sarcoma, lymphoma, leukemia, melanoma, carcinoma, glioblastoma and
CC astromytoma, with defective tumour suppressor genes, oncogenes,
CC infectious diseases, allergy or autoimmune diseases. The present
CC sequence is a targeting peptide termed RGD-4C. This may be incorporated
CC into the amino terminal region of a KDEL receptor inhibitor protein
CC downstream from a cleavably removed sequence to improve its activity or
CC alter its immunogenicity.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 14
AAY54271
ID AAY54271 standard; Peptide; 9 AA.
XX
AC AAY54271;
XX
DT 06-APR-2000 (first entry)
XX
DE Alpha Vbeta-3 binding peptide sequence.
XX
KW Envelope protein; mutant; retrovirus; surface protein shedding;
KW envelope protein stability; gene therapy; drug therapy; cancer;
KW adenosine deaminase deficiency; thalassemia; hemophilia; diabetes;
KW alpha-anti trypsin deficiency; brain disorder; neural disorder;
KW phenylketonuria; growth disorder; heart disease; immune disease.
XX
OS Unidentified.
XX
PN WO9960110-A2.
XX
PD 25-NOV-1999.
XX
PF 20-MAY-1999; 99WO-US11155.
XX
PR 20-MAY-1998; 98US-0086149.
XX
PA (UYTE-) UNIV TENNESSEE RES CORP.
XX
PI Albritton LM, Zavorotinskaya T;

XX
DR WPI; 2000-116313/10.
XX
PT Novel isolated nucleic acid, useful for gene therapy -
XX
PS Example 10; Page 84; 190pp; English.
XX
CC The specification describes mutant retrovirus envelope proteins. The
CC envelope protein coding sequence can be mutated to encode a mutant
CC envelope protein with a substitution of one or more amino acids in at
CC least one motif of the retrovirus protein. The mutant protein fragment
CC allows for decreased shedding of the surface protein by suppressing
CC precursor cleavage and increase envelope stability and fusion of
CC retroviruses with cell membranes, while maintaining mutant envelope
CC protein incorporation into a virion, and viral titers of about two orders
CC of magnitude within that observed for wild-type retrovirus when the
CC protein or fragment is expressed on the surface of a retroviral particle.
CC The proteins have an increased ability to penetrate targets, typically
CC cells and a correspondingly increased ability to deliver nucleic acids or
CC drugs. The mutated nucleic acid is useful for gene and drug therapy,
CC especially as drug delivery vehicles. The retrovirus particles can be
CC utilized to transduce eukaryotic cells. The transduced cells are useful
CC in the treatment of cancer in a human. Other diseases contemplated for
CC treatment include adenosine deaminase deficiency (ADA), thalassemia,
CC hemophilia, diabetes, alpha-anti trypsin deficiency, brain and neural
CC disorders, phenylketonuria, growth disorders, heart diseases and immune
CC diseases. The present sequence was used in the course of the invention,
CC to quantitate targeted retroviral vector gene delivery in vivo.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 15
AAE11044
ID AAE11044 standard; peptide; 9 AA.
XX
AC AAE11044;
XX
DT 18-DEC-2001 (first entry)
XX
DE RGD-containing peptide.
XX
KW Tumour necrosis factor; TNF; cytokine; cytostatic; virucide;
KW TNF related apoptosis inducing ligand; TRAIL; cancer; viral infection;
KW human immunodeficiency virus; HIV; leukaemia; gene therapy; lymphoma;
KW melanoma.
XX
OS Unidentified.
XX
PN US6284236-B1.
XX
PD 04-SEP-2001.
XX
PF 26-MAY-1999; 99US-0320424.
XX
PR 29-JUN-1995; 95US-0496632.
PR 01-NOV-1995; 95US-0548368.
PR 25-JUN-1996; 96US-0670354.
PR 26-MAR-1998; 98US-0048641.
PR 10-NOV-1998; 98US-0190046.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Wiley SR, Goodwin RG;

XX
DR WPI; 2001-595463/67.
XX
PT New tumor necrosis factor related apoptosis inducing ligand
PT polypeptides for treating viral infections (e.g. bovine viral diarrhea
PT or human immunodeficiency virus), or cancers (e.g. leukemia or
PT lymphoma) -
XX
PS Disclosure; Column 11; 41pp; English.
XX
CC The invention relates to a cytokine designated as tumour necrosis
CC factor (TNF) related apoptosis inducing ligand (TRAIL), which induces
CC apoptosis of certain target cells, including cancer cells and virally
CC infected cells. The TRAIL polypeptides are useful in killing cancer
CC cells, in treating viral infections (e.g. bovine viral diarrhoea or
CC human immunodeficiency virus (HIV)) and cancers (e.g. leukaemia,
CC lymphoma and melanoma), as a research reagent useful in studying
CC apoptosis including the regulation of programmed cell death. TRAIL
CC DNA sequences may be employed in developing a gene therapy approach
CC to treating disorders mediated by defective or insufficient amounts
CC of TRAIL, in the production of TRAIL polypeptides and as probes or
CC primers in polymerase chain reactions (PCR). The present sequence is
CC a RGD-containing peptide that binds an integrin associated with
CC tumour. This sequence is used to construct a fusion protein
CC comprising TRAIL protein.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
|||||||||
Db 1 cdcrgdcfc 9

RESULT 16
AAE06279
ID AAE06279 standard; peptide; 9 AA.
XX
AC AAE06279;
XX
DT 25-SEP-2001 (first entry)
XX
DE Tumour homing peptide used for homing pro-apoptotic conjugates.
XX
KW Chimeric prostate-homing pro-apoptotic peptide; prostate-homing peptide;
KW antimicrobial peptide; prostate cancer; tumour homing molecule;
KW cytostatic; RGD motif.
XX
OS Synthetic.
XX
PN WO200153342-A1.
XX
PD 26-JUL-2001.
XX
PF 16-JAN-2001; 2001WO-US01362.
XX
PR 21-JAN-2000; 2000US-0489582.
XX
PA (BURN-) BURNHAM INST.
XX
PI Ruoslahti EI, Pasqualini R, Arap W, Bredesen DE, Ellerby HM;
XX
DR WPI; 2001-451901/48.
XX
PT Novel chimeric prostate-homing pro-apoptotic peptide, used to treat
PT prostate cancer, comprises a prostate-homing peptide linked to an
PT antimicrobial peptide -
XX
PS Example 3B; Page 84; 176pp; English.

```
XX
CC   The patent discloses novel chimeric prostate-homing pro-apoptotic
CC   peptide which comprises a prostate-homing peptide linked to an
CC   antimicrobial peptide, where the chimeric peptide is selectively
CC   internalised by and exhibits high toxicity to prostate tissue and
CC   where the antimicrobial peptide has low mammalian cell toxicity when
CC   not linked to prostate-homing peptide. The chimeric peptide is used
CC   to direct an antimicrobial peptide in vivo to a prostate cancer, to
CC   induce selective toxicity in vivo in a prostate cancer, and to treat
CC   a patient with prostate cancer. The present peptide sequence is a
CC   tumour homing molecule containing a RGD motif. This sequence is
CC   useful in the homing of pro-apoptotic conjugates of the invention.
XX
SQ   Sequence   9 AA;


  Query Match             100.0%;  Score 65;  DB 22;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CDCRGDCFC 9
          |||||||||
Db      1 cdcrgdcfc 9


RESULT  17
AAB97086
ID   AAB97086 standard; peptide; 9 AA.
XX
AC   AAB97086;
XX
DT   02-AUG-2001  (first entry)
XX
DE   Integrin-binding peptide #4.
XX
KW   Integrin; avB3; avB5; analgesic; cytostatic; macrocyclic chelant;
KW   metal chelate formation; metalloradiopharmaceutical;
KW   magnetic resonance imaging; MRI; disease diagnosis;
KW   systemic radiotherapy; bone pain; bone cancer; antagonist.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The amino group of the residue at position 1
FT                          forms a peptide bond with the carboxy group of
FT                          the residue at position 9 to form a cyclic
FT                          molecule"
FT   Modified-site   9
FT                   /note= "The amino group of the residue at position 1
FT                          forms a peptide bond with the carboxy group of
FT                          the residue at position 9 to form a cyclic
FT                          molecule"
XX
PN   WO200119838-A1.
XX
PD   22-MAR-2001.
XX
PF   07-SEP-2000; 2000WO-US24482.
XX
PR   13-SEP-1999;   99US-0153512.
XX
PA   (DUPO ) DU PONT PHARM CO.
XX
PI   Liu S;
XX
DR   WPI; 2001-389600/41.
XX
PT   New nitrogen containing macrocyclic chelant compounds used in metal
PT   chelates for e.g. X-ray imaging and for attaching diagnostic and
PT   therapeutic isotopes to biologically active targeting molecules -
XX
PS   Disclosure; Page 72; 121pp; English.
XX
CC   The present sequence is provided in a specification relating to novel
CC   nitrogen containing macrocyclic chelant compounds. The compounds are
CC   used for forming metal chelates used as diagnostic or therapeutic
CC   metalloradiopharmaceuticals, or magnetic resonance imaging (MRI)
CC   contrast agents. They are also used for attaching metal ions to
CC   bio-directing groups including proteins, peptides, peptidomimetics
CC   and non peptides that bind in vivo to a receptor or enzyme that is
CC   expressed or up-regulated at a site or in a disease state. The
CC   metallopharmaceuticals are used in disease diagnosis by MRI or in
CC   treating disease by systemic radiotherapy. Radiolanthide chelates
CC   with phosphonomethyl and optionally carboxymethyl groups on the four
CC   N atoms can be used for treating bone pain and bone metastases.
CC   The macrocyclic chelants rapidly form stable metal chelates. The
CC   present sequence binds with high affinity to the integrins avB3 and
CC   avB5.
XX
SQ   Sequence   9 AA;


  Query Match             100.0%;  Score 65;  DB 22;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 6.4e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CDCRGDCFC 9
          |||||||||
Db      1 cdcrgdcfc 9


RESULT  18
AAB20271
ID   AAB20271 standard; Peptide; 9 AA.
XX
AC   AAB20271;
XX
DT   14-MAY-2001  (first entry)
XX
DE   Peptide that specifically targets tumour blood vessels.
XX
KW   Tumour; breast carcinoma; Karposi's sarcoma; melanoma;
KW   fiberless radiative effector; therapy; imaging.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 4..6
FT                   /note= "RGD motif"
XX
PN   WO200108660-A2.
XX
PD   08-FEB-2001.
XX
PF   26-JUL-2000; 2000WO-US20292.
XX
PR   02-AUG-1999;   99US-0366314.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Philbert MA,  Tjalkens R,  Aylott JW,  Clark HA,  Monson EE;
PI   Kopelman R;
XX
DR   WPI; 2001-182851/18.
XX
PT   Composition for destroying or inhibiting growth of tumour cells and
PT   for imaging tumours or other biological targets, has molecular
PT   recognition element attached to fiberless radiative effector having
PT   a toxic agent -
XX
PS   Disclosure; Page 35; 95pp; English.
XX
CC   The present sequence is that of a peptide that specifically binds
CC   to tumour blood vessels.  It includes an RGD motif.  The peptide,
CC   and conjugates containing it, selectively binds to various tumours,
```

CC including breast carcinomas, Karposi's sarcoma and melanoma. The
CC peptide can be used as the molecular recognition element of novel
CC fiberless radiative effectors (FREs) of the invention. The
CC invention is related to cell or pathogen destruction via FREs
CC that encapsulate a radical generator. The FREs include a polymer
CC matrix, a photodynamic or radiodynamic dye which produces free
CC radicals upon stimulation, cloaking material, and at least 1
CC molecular recognition element for targeting to a biological target,
CC e.g. the present peptide. They are useful in various in vitro and
CC in vivo procedures, destroying or inhibiting the growth of
CC biological targets (pathogens, macromolecules, tumour cells in
CC culture or in the body), in therapies including chemotherapy,
CC radiation therapy, antibiotic and vaccine therapy.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
||||||||| |
Db 1 cdcrgdcfc 9

RESULT 19
AAB50242
ID AAB50242 standard; peptide; 9 AA.
XX
AC AAB50242;
XX
DT 13-MAR-2001 (first entry)
XX
DE Enhanced infectivity adenoviral vector fibre replacement ligand.
XX
KW Adenoviral vector; gene therapy; infectability;
KW tumour-specific replication.
XX
OS Unidentified.
XX
PN WO200067576-A1.
XX
PD 16-NOV-2000.
XX
PF 12-MAY-2000; 2000WO-US13114.
XX
PR 12-MAY-1999; 99US-0133634.
XX
(UABR-) UAB RES FOUND.
XX
PI Curiel DT, Krasnykh VN, Alemany R, Dmitriev I;
XX
DR WPI; 2001-122702/13.
XX
PT New infectivity-enhanced, conditionally-replicative adenovirus
PT containing a modified wild type adenoviral fiber, useful for cancer
PT therapy -
XX
PS Claim 8; Page 70; 104pp; English.
XX
CC The present invention provides an adenoviral vector with an enhanced
CC ability to infect tumour cells and which is conditionally replicative,
CC enabling replication in only one cell type. This can be used in the
CC gene therapy treatment of cancers.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDCRGDCFC 9
||||||||| |
Db 1 cdcrgdcfc 9

Search completed: May 29, 2002, 10:42:20
Job time: 38 sec